DOCUMENT RESUME

ED 211 413 SO 013 779

TITLE	Procedural Handbook, 1978-79 Music Assessment.
INSTITUTION	Education Commission of the States, Denver, Colo. National Assessment of Educational Progress.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
REPORT NO	ISBN-0-89398-188-5; NAEP-10-MU-40
PUB DATE	Dec 81
GRANT	NIE-G-80-0003
NOTE	83p.; For a related document, see SO 013 741.
AVAILABLE FROM	National Assessment of Educational Progress, 1860 Lincoln St., Suite 700, Denver, CO 80295 ($7.70).
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Administration; Cognitive Ability; Data Collection; *Educational Assessment; Elementary Secondary Education; *Evaluation Methods; *Music; Music Education; National Surveys; Sampling; Scoring; Student Attitudes
IDENTIFIERS	*National Assessment of Educational Progress; Second Music Assessment (1979)

ABSTRACT

This handbook describes the procedures used to administer and analyze the results of the 1978-79 national music assessment of the National Assessment of Educational Progress (NAEP). It also describes changes in procedures between the first NAEP music assessment conducted in 1971-72 and the 1978-79 assessment. Each assessment surveyed the music achievement and attitudes of American 9-, 13- and 17-year-olds, using a deeply stratified, multistage probability sample design. The purpose of the handbook is to provide detailed procedural information for people interested in replicating the assessment or in need of more information than is provided in the reports containing assessment data. The eight chapters cover objectives redevelopment, exercise creation, preparation of assessment booklets, sampling, data collection, scoring, data analysis, and reporting. Each chapter explains the basic procedures used for the 1978-79 assessment and contrasts these procedures to those used in earlier years if there were changes. Appendices, which comprise over half of the handbook, contain materials covering definitions of reporting groups, forms used to gather background information about students and schools, response rates, computation of achievement measures and procedures for smoothing respondent weights. A glossary of NAEP terms is provided at the end of the book. Primary type of information provided by report: Procedures (Overview). (Author/RM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211413

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PROCEDURAL HANDBOOK
1978-79 MUSIC ASSESSMENT

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Pratt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Report No. 10-MU-40

by the
National Assessment of Educational Progress

Education Commission of the States
Suite 700, 1860 Lincoln Street
Denver, Colorado 80295

SP013779

December 1981

The National Assessment of Educational Progress is funded by the National Institute of Education under a grant with the Education Commission of the States. It is the policy of the Education Commission of the States to take affirmative action to prevent discrimination in its policies, programs and employment practices.

Library of Congress
Catalog Card Number: 72-183665

Although a few early National Assessment reports have individual catalog card numbers, all recent reports have been assigned the above series number.

ISBN 0-89398-188-5

The National Assessment of Educational Progress is an education research project mandated by Congress to collect and report data, over time, on the performance of young Americans in various learning areas. National Assessment makes available information on assessment procedures and materials to state and local education agencies and others.

The work upon which this publication is based was performed pursuant to Contract No. OEC-0-74-0506 of the National Center for Education Statistics and the National Institute of Education; also, Grant No. NIE-G-80-0003 of the National Institute of Education. It does not, however, necessarily reflect the views of those agencies.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
Education Commission of the States

Robert D. Ray, Governor of Iowa, Chairperson,
Education Commission of the States
Robert C. Andringa, Executive Director,
Education Commission of the States
Roy H. Forbes, Director, National Assessment

All National Assessment reports and publications are available through NAEP offices at the address shown at the bottom. Reports ordered from NAEP are processed within three days.

Some of the more recent reports are also available from the Superintendent of Documents, and most all are available from ERIC. Write to the National Assessment office for a current publications list.

National Assessment reports related to this report:

MUSIC

1st Assessment (1971-72)

03-MU-01	The First National Assessment of Musical Performance, February 1974	$ 1.00
03-MU-02	A Perspective on the First Music Assessment, April 1974	1.00
03-MU-03	An Assessment of Attitudes Toward Music, September 1974	1.10
03-MU-00	The First Music Assessment: An Overview, August 1974	1.00
03-MU-20	Music Technical Report: Exercise Volume, December 1975	25.00
03-MU-21	Music Technical Report: Summary Volume, November 1975	4.40

NOTE: A cassette supplementing the music reports including musical stimuli and actual performance by 9-, 13- and 17-year olds is available for $1.00.

2nd Assessment (1978-79)

10-MU-01	Music 1971-79: Results From the Second National Music Assessment, November 1981	7.00
10-MU-25	The Second Assessment of Music, 1978-79, Released Exercise Set, April 1980	15.20
10-MU-40	Procedural Handbook: 1978-79 Music Assessment, December 1981	7.70

NOTE: A cassette tape with music stimuli for released exercises supplementing Report no. 10-MU-25 is now available for $2.00.

(Continued on Inside Back Cover)

TABLE OF CONTENTS

LIST OF TABLES v

FOREWORD vii

ACKNOWLEDGMENTS ix

INTRODUCTION xi

CHAPTER 1 Objectives Redevelopment 1

CHAPTER 2 Development of Exercises 5
Exercise Development 5
Field Tryouts 5
Exercise Reviews and Final Selection 6

CHAPTER 3 Preparation of Assessment Materials 7
Overview of Packaging 7
Differences in Packaging: 1971-72 and 1978-79 . . . 8

CHAPTER 4 Sampling 9
Overview of the National Assessment Sample Design 10
Survey Weights 10

CHAPTER 5 Data Collection 13

CHAPTER 6 Scoring 17

CHAPTER 7 Data Analysis 19
Measures of Achievement 19
Estimating Variability in Achievement Measures . . . 21
Controlling Nonrandom Errors 22

CHAPTER 8 Reports About the Music Assessment 23
Reports 23
Public-Use Data Tape and User Services 24

APPENDIX A Definitions of National Assessment Reporting Groups 25

APPENDIX B Forms Used To Obtain Background Information 29

APPENDIX C Response Rates for Assessment Samples 53

APPENDIX D Computation of Measures of Achievement, Changes in Achievement and Standard Errors 55
Measures of Achievement 55
Computation of Standard Errors 57

APPENDIX E Adjustment of Respondent Weights by Smoothing To Reduce Random Variability of Estimated Population Proportions 61
Background 61
Smoothing Procedures Used by National Assessment 62
The Current Smoothing Procedure. 63
Adjustment of Weights by Users 64
Changes in Smoothed Proportions as New Assessments Are Completed 66

GLOSSARY OF NATIONAL ASSESSMENT TERMS 67

BIBLIOGRAPHY 73

LIST OF TABLES

TABLE 1. Number of PSUs and Schools Within PSUs Selected in 1971-72 and 1978-79 11

TABLE 2. School Cooperation Rates, 1978-79 Assessment 13

TABLE 3. Number of Students Responding to Each Item Booklet in 1978-79 Assessment, by Age 14

TABLE 4. Quality-Control Summaries for the 1978-79 Music Assessment Open-Ended Exercises, by Age 18

TABLE C-1. Number of Students Assessed and Percent of Sample Covered, by Age and Assessment Year 53

TABLE D-1. Average Number of Respondents in Reporting Groups Taking an Item Booklet, by Age and Assessment Year 59

TABLE D-2. Estimated Current Population Proportions of National Assessment Reporting Groups for In-School Students 60

TABLE E-1. Smoothing Cells Used for the 1978-79 Smoothing Procedures 63

TABLE E-2. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 9-Year-Olds 64

TABLE E-3. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 13-Year-Olds 65

TABLE E-4. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 17-Year-Olds 66

FOREWORD

When the U.S. Office of Education was chartered in 1867, one charge to its commissioners was to determine the nation's progress in education. The National Assessment of Educational Progress (NAEP) was initiated a century later to address, in a systematic way, that charge.

Since 1969, the National Assessment has gathered information about levels of educational achievement across the country and reported its findings to the nation. It has surveyed the attainments of 9-year-olds, 13-year-olds, 17-year-olds and sometimes adults in art, career and occupational development, citizenship, literature, mathematics, music, reading, science, social studies and writing. All areas have been periodically reassessed in order to detect any important changes. To date, National Assessment has interviewed and tested nearly 1,000,000 young Americans.

Learning-area assessments evolve from a consensus process. Each assessment is the product of several years of work by a great many educators, scholars and lay persons from all over the nation. Initially, these people design objectives for each subject area, proposing general goals they feel Americans should be achieving in the course of their education. After careful reviews, these objectives are given to exercise (item) writers, whose task it is to create measurement instruments appropriate to the objectives.

When the exercises have passed extensive reviews by subject-matter specialists, measurement experts and lay persons, they are administered to probability samples. The people who compose these samples are chosen in such a way that the results of their assessment can be generalized to an entire national population. That is, on the basis of the performance of about 2,500 9-year-olds on a given exercise, we can make generalizations about the probable performance of all 9-year-olds in the nation.

After assessment data have been collected, scored and analyzed, the National Assessment publishes reports and disseminates the results as widely as possible. Not all exercises are released for publication. Because NAEP will readminister some of the same exercises in the future to determine whether the performance levels of Americans have increased, remained stable or decreased,

it is essential that they not be released in order to preserve the integrity of the study.

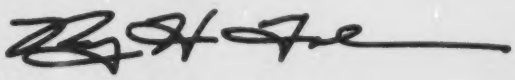
ACKNOWLEDGMENTS

Many organizations and individuals have made substantial contributions to the music assessments. Not the least of those to be gratefully acknowledged are the administrators, teachers and students who cooperated so generously during the collection of the data.

Special acknowledgment must go to the many music educators and specialists who provided their expertise in the development, review and selection of the assessment objectives and exercises.

Administration of the music assessment was conducted by the Research Triangle Institute, Raleigh, North Carolina. Scoring and processing were carried out by Westinghouse DataScore Systems, Iowa City, Iowa, and by the National Assessment staff.

Every assessment is the result of a collaborative effort by the National Assessment staff. Many persons contributed to the music assessments. Special thanks are extended to Jan Pearson for development, scoring and analysis; Donald T. Searls and Eugene Johnson for sampling; Dunlap Scott for coordinating data collection; Gwen Edwards for data processing support; Marci Reser and Deborah Houy for report production; and Barbara Holmes for editorial supervision.



Roy H. Forbes
Director

INTRODUCTION

The National Assessment of Educational Progress (NAEP) has completed two assessments of music, the first conducted in 1971-72 and the second during 1978-79. Each assessment surveyed the music achievement and attitudes of American 9-, 13- and 17-year-olds, using a deeply stratified, multistage probability sample design. This handbook describes the procedures used to administer and analyze the results of the 1978-79 music assessment. It also describes changes in procedures between the assessments.

To measure changes in performance between 1971-72 and 1978-79, approximately half of the exercises in the first assessment were reassessed in the second under almost identical administrative conditions. To measure the status of music achievement in 1978-79, National Assessment consultants revised the objectives used in the first assessment and developed additional exercises to provide coverage of the revised objectives.[1]

Approximately 20,000 9-year-olds, 25,000 13-year-olds and 22,000 17-year-olds participated in the 1978-79 music assessment. Because National Assessment reports results for groups of students, not individuals, it is not necessary for each student to respond to every item (exercise).[2] Each respondent completed only one item booklet of about 45 minutes in length. Between 2,400 and 2,800 students responded to each booklet. In 1978-79 there were eight exercise booklets for 9-year-olds, nine booklets for 13-year-olds and eight booklets for 17-year-olds.

The exercises for each assessment were administered by a professional data collection staff to minimize the burden on participating schools and to maximize uniformity of assessment conditions. Instructions and items were recorded on a paced audio tape and played back to students to reduce the potential effect of reading difficulties and to

---

[1] Because of cuts in funding, exercises assessing music performance were not administered in the 1978-79 assessment.

[2] National Assessment uses the term "exercise" to mean an assessment item. The terms "exercise" and "item" are used interchangeably in this report.

ensure that all students moved through the packages at the same speed.

The majority of the items were multiple-choice; a few exercises from the first assessment that were reassessed in 1978-79 were open-ended. Each item included one or more parts.

Multiple-choice items were scored by an optical scanning machine; open-ended items were hand-scored by trained scorers using scoring guides developed in the 1971-72 assessment. These scoring guides defined categories of acceptable and unacceptable responses.

National Assessment reports estimated percentages of correct responses for single items. When a report indicates that "85% of the 17-year-olds gave a correct response," it means that an estimated 85% of the 17-year-olds would have given a correct response if all the 17-year-olds in schools across the country had been assessed. In addition to reporting national results, NAEP provides data on the performance of various population subgroups within the national population, defined by sex, race/ethnicity, region of the country, size and type of community lived in and level of parental education. For the 1978-79 music assessment, some data are available on the extent of the music training and/or experience, either formal or informal, of the 9-, 13- and 17-year-olds. National Assessment aggregates percentages of success on various sets of items to provide data on changes in performance between assessments and on the differential performance of population subgroups.

This handbook describes the procedures used to develop, administer and analyze the results of the 1978-79 music assessment. Its primary purpose is to provide detailed procedural information for people interested in replicating the assessment or in need of more information than is provided in the reports containing assessment data. The eight chapters cover objectives redevelopment, exercise creation, preparation of assessment booklets, sampling, data collection, scoring, data analysis and reporting. Each chapter explains the basic procedures used for the 1978-79 assessment and contrasts these procedures to those used in earlier years (if there were changes).

Appendicular materials cover definitions of reporting groups, forms used to gather background information about students and schools, response rates, computation of achievement measures and procedures for smoothing respondent weights. A glossary of National Assessment terms is provided at the end of the book.

CHAPTER 1

OBJECTIVES REDEVELOPMENT

The primary goal of the National Assessment of Educational Progress (NAEP) is to report on the current educational status of young Americans and to monitor any changes in achievement over time. For each learning area to be assessed, NAEP asks consultants to develop objectives that define the subject area. Since the objectives provide guidelines for exercise developers, consultants are asked to include examples of the knowledge, skills and attitudes to be assessed at each age level.

Education in America is a collaborative enterprise involving a great many people with widely differing philosophies. Providing information about education nationwide would be considerably easier if there were a consensus about the means and ends of American education, but the fact is that Americans have conflicting and sometimes contradictory values regarding the goals of education and the means for achieving those goals. To develop an assessment that is truly national in scope and takes into account the diversity of curricula, values and goals across the country, National Assessment employs a consensus process for developing objectives, involving representation of many different groups of people.

Several types of consultants helped develop the 1978-79 music objectives. College and university specialists in music ensured that the objectives included important concepts that the schools should be teaching. Educators, including classroom teachers, curriculum supervisors and persons involved in teacher education, made sure that the objectives included concepts, skills and attitudes that the schools are presently teaching, as well as those they should be teaching. Concerned citizens, parents and other interested lay persons had to agree that the objectives were important for young people to achieve, were free of educational jargon and were not biased or offensive to any groups. Consultants were representative of the different regions of the country and minority groups. They also represented a range of experience with students of different ages and community types.

The objectives for the 1978-79 music assessment were developed during 1972-73. (The second assessment of music was

originally scheduled to occur earlier than 1978-79.) This redevelopment was done through conferences organized and conducted by the National Assessment staff. The objectives redevelopment and review conferences involved many music educators and music specialists. The objectives also were reviewed by a panel of lay persons to ensure that they covered topics generally considered to be important and that they were not biased.

The objectives for the second music assessment were built upon the framework of the objectives used for the first assessment, with more emphasis placed on affective responsiveness. Consultants stressed the fact that some of the most important goals in music are difficult to state in observable terms and are difficult to measure. One such goal is to develop aesthetic sensitivity within each student. Another is to help the student develop his or her music potential as far as possible during elementary and secondary school years. These goals are as important for the nonperformer as for the performer.

Following the redevelopment of the objectives, another large group of music consultants reviewed the objectives and assigned weights to define how much coverage each objective should have at each age.

Prior to the beginning of development of the music exercises, an advisory group of music consultants was brought together to give direction and advice to the NAEP staff on further refinement of the objectives and development of the assessment.

Following is the outline of the objectives used for the 1978-79 music assessment (Music Objectives, Second Assessment, 1980).

I. Value music as an important realm of human experience
   - A. Be affectively responsive to music
   - B. Be acquainted with music from different nations, cultures, periods, genres and ethnic groups
   - C. Value music in the life of the individual, family and community
   - D. Make and support aesthetic judgments about music

II. Perform music
   - A. Sing (without score)
   - B. Play (without score)
   - C. Sing or play from a written score
   - D. Play or sing a previously prepared piece

III. Create music
   - A. Improvise
   - B. Represent music symbolically

IV. Identify the elements and expressive controls of music
   - A. Identify the elements of music
   - B. Identify the relationships of ele-

ments in a given composition

C. Demonstrate an understanding of a variety of musical terms, expression markings and conducting gestures in a musical context

V. Identify and classify music historically and culturally

A. Identify and describe the features that characterize a variety of folk, ethnic, popular and art music

B. Identify and describe the music and musical style of the various stylistic periods in Western civilization (e.g., Medieval, Renaissance, Baroque, Classical, Romantic). Identify representative composers of each period

C. Cite examples of ways in which man utilizes music in his social and cultural life

CHAPTER 2

DEVELOPMENT OF EXERCISES

Exercise Development

Groups of consultants and advisors weighted the relative importance of each objective to determine how much testing time each should have. Exercises reassessed in 1978-79 were reclassified to match the 1978-79 objectives so that change and status reporting would be similar.

Exercises were developed in groups to measure a single objective or portion of an objective. Every effort was made to develop groups of exercises that could be given to all of the age levels, with changes in vocabulary, stimuli and format when necessary.

Item development took place from February 1977 to January 1978, with approximately 50 music specialists and music educators participating as item writers. Because of reduced funding, which prevented the assessment of music performance, development was directed to the three remaining objectives: Objectives I, IV and V. (Development of music performance exercises was conducted during fall 1978 but, as previously mentioned, they were not included in the 1978-79 assessment.)

There were many exercise writing conferences, review conferences, two field pretestings and a final selection conference. The methods for achieving quality exercises varied with each conference. Typically, each group of music consultants worked on developing exercises for a specific objective or for a specific age. When a large group met, the consultants divided into subgroups for this purpose. Individuals or small groups generated items. Then the entire group came together to review all new items. During item review, consultants considered age-level appropriateness, accuracy of content, how well the item measured an objective or subobjective and readability. Following the review procedure, more writing occurred, followed by more reviewing, and so on until the group had completed the assignment. Exercises that came out of the development conferences were edited by the National Assessment staff to fit NAEP format and technical requirements and were prepared for the field "tryouts," or pretesting.

Field Tryouts

Music exercises for the 1978-

79 music assessment were field tested in schools across the country during spring 1977 and again in fall 1977. The exercises were field tested in situations as close as possible to those of the actual assessment to discover potential problems in wording, directions or administrative procedures and to collect item statistics, timing information and scoring information."Tryout" schools were selected to represent high- and low-income communities as well as more typical communities. The tryouts were conducted with students in at least four classrooms (approximately 100 students) at each of the ages assessed. In order to simulate actual assessment field procedures, students recorded their answers in the test booklets; directions and questions were read to students from an audio tape; and National Assessment staff members, rather than classroom teachers, administered the test.

Completed tryout booklets were scored and data were analyzed. Data for the items, as well as the administrators' reports of any field problems, helped both NAEP staff and consultants evaluate and revise the exercises. Following the spring 1977 field tests, a development process like that described above was used to produce additional exercises or to revise or rework exercises from the first tryouts in preparation for fall 1977 field testing.

Exercise Reviews and Final Selection

Each music exercise that had been pretested was included in a pool of items that was reviewed in a series of conferences by at least 25 different consultants: music educators, music specialists, classroom teachers and interested lay citizens. Exercises for each age group were reviewed for appropriateness by teachers who taught students at that age. Lay citizens, representing a variety of occupations and interests, also reviewed the exercises, checking for sex or race bias and considering the general importance of each exercise.

After the review conferences, a final set of exercises for the 1978-79 assessment was selected by another group of music consultants.

CHAPTER 3

PREPARATION OF ASSESSMENT MATERIALS

Overview of Packaging

Following the selection of music exercises, National Assessment staff grouped and sequenced them into exercise booklets. The 1978-79 assessment was a combined assessment of music, art and writing. Because of the length of many of the art and writing exercises and because of the many handouts used with them, only two areas were packaged together to avoid administrative problems. Booklets included either music and writing exercises or art and writing exercises. Since different ages received somewhat different sets of exercises, there were separate booklets for each age level. Thus, exercises for 9-year-olds were not sequenced in the same order as those for 13-year-olds, and so forth.

The following constraints were observed in preparing exercise booklets:

1. Each booklet contained exercises of varying difficulty so that students would not become bored by many easy exercises or discouraged by many difficult exercises.

2. Exercises could not cue other exercises. In other words, the answer to one exercise could not be contained in another exercise in the same booklet.

3. The booklets included writing and music exercises. Each booklet was timed so that it would take no more than 45 minutes of a student's time -- the length of a typical class period. Booklets contained approximately 30-35 minutes of exercise time and an additional 10-15 minutes of introductory material, instructions and background questions. The total amount of exercise time devoted to music across <u>all</u> packages for an age level was 77 minutes at age 9, 81 minutes at age 13 and 86 minutes at age 17.

4. Booklets were designed to be, insofar as possible, parallel with respect to the number of different objectives measured and difficulty levels. Items measuring a particular objective were scattered throughout the booklets so that many different students would respond to questions related to a particular objective.

National Assessment makes every effort to minimize difficulties connected with the testing situation so that results will be, as nearly as possible, an accurate reflection of what students know and can do. For example, students marked their answers directly in the assessment booklets, not on separate answer sheets. It was felt that this procedure would reduce possibilities for errors in marking answer sheets, especially for the younger students. To minimize guessing, students were encouraged to select the "I don't know" response option included with each multiple-choice item or to write "I don't know" on the answer line for open-ended questions if they felt they did not know the answer to a question.

Paced audio tapes were prepared for each exercise booklet. Instructions, most of the written portions of an exercise stimulus and response options were read aloud to minimize the effect of any reading difficulties and to ensure that all students moved through the booklets at the same speed. During the field testing of the exercises, administrators had determined the time needed for most students to respond to an item. In addition, the use of tapes helped to ensure uniform assessment conditions across the country.

Differences in Packaging: 1971-72 and 1978-79

National Assessment makes every effort to make assessment conditions for items measuring change identical from assessment to assessment so that any changes observed will be attributable to changes in achievement rather than a response to an altered testing condition.

However, in 1971-72, music and social studies were assessed together. In 1978-79, music and writing were assessed together. Because many of the writing exercises were quite lengthy, the number of exercises per package in 1978-79 was considerably fewer than the number of exercises per package in 1971-72.

Some modification occurred in the tape recorded reading of the exercises from 1971-72 to 1978-79. In 1971-72, the "I don't know" choice was read aloud for each exercise. In 1978-79, this response choice was read aloud only for those 1971-72 exercises reassessed to measure changes in achievement.

CHAPTER 4

SAMPLING

This chapter gives an overview of the procedures used in designing and selecting the National Assessment samples for the 1971-72 and 1978-79 music assessments.[1] Sample design and selection for both assessments were conducted by the Research Triangle Institute, Raleigh, North Carolina, and monitored by National Assessment staff.

The target populations for each of the assessments included 9-, 13- and 17-year-olds[2] enrolled in either public or private schools at the time of the assessment. Seventeen-year-olds who either left school before graduating or graduated early and adults 26-35 years old were also included in the 1971-72 assessment. However, assessment of out-of-school 17-year-olds and adults is quite expensive, and in 1978-79 funds to assess these populations were not available.

Age-eligible persons who were functionally handicapped to the extent that they could not participate in an assessment were not considered part of the target populations. Specific groups excluded were: non-English-speaking persons, those identified as non-readers, persons physically or mentally unable to respond and persons in institutions or attending schools established for the physically or mentally handicapped.

National Assessment does not follow up specific individuals from one assessment to the next. In other words, the students who participated in the 1978-79 assessment were not the same ones who participatd in 1971-72. However, in each assessment year, participants are carefully selected to represent each age level. For example, National Assessment assessed one set of probability samples of 9-year-olds in 1972 and a totally different set of probability samples of

---

[1] For detailed information about the 1978-79 National Assessment sampling procedures, see Final Report...Sampling and Weighting Activities for Assessment Year 10 (1980).

[2] Definition of 1978-79 assessment age groups are: 9-year-olds -- born during calendar year 1969; 13-year-olds -- born during calendar year 1965; and 17-year-olds -- born October 1, 1961, through September 30, 1962.

9-year-olds in 1979. Thus, when we say that 9-year-olds' achievement declined between 1972 and 1979, we mean that students who were 9 years old in 1972 correctly answered the same questions more often than those who were 9 years old in 1979.

The National Assessment samples were designed to provide approximately 2,500 respondents per exercise. These numbers allow reporting of data for the nation and for the subgroups defined in Appendix A. The sample design used to obtain representative samples of the target populations was modified somewhat between assessments. However, these minor changes did not affect the comparability of the samples. An overview of the general sample design approach follows.

Overview of the National Assessment Sample Design

National Assessment used a deeply stratified, three-stage national probability sample design with oversampling of low-income and rural areas. In the first stage, the United States was divided into geographical units of counties or groups of contiguous counties meeting a minimum population size requirement. These units, called primary sampling units (PSUs), were stratified by region and size of community. From the list of PSUs, a sample of PSUs was drawn, without replacement, with probability proportional to population size measures, representing all regions and sizes of communities. Oversampling of low-income and extreme-rural areas was first performed at this stage by adjusting the estimated population size measures of those areas to increase sampling rates. Within PSUs, Census Employment Survey Data were used to further delineate and oversample low-income areas. Counties with high proportions of rural families were also oversampled.

In the second stage, all public and private schools within each PSU selected in the first stage were listed. Schools within each PSU were selected, without replacement, with probabilities proportional to the number of age-eligibles in the school.

The third stage of sampling occurred during the data collection period. A list was made of all age-eligible students within each selected school. A simple random selection of eligible students, without replacement, was obtained, and item booklets were administered to selected students. Specially trained personnel selected the sample and administered the booklets.

Survey Weights

Not all respondents in the sample had the same probability of selection because some subpopulations were oversampled, and adjustments were made to compensate for some schools' refusals to participate and for student nonre-

sponse. The selection probability for each individual was computed, and its reciprocal was used to weight each response in any statistical calculation to compensate for unequal rates of sampling and to ensure proper representation in the population structure. Procedures used to assign weights are discussed in Chapter 7 and Appendix E.

The number of PSUs, schools within PSUs and students within schools was determined by optimum sampling principles. That is, a sample design was selected that would achieve the maximum precision for a given level of resources. Table 1 displays the number of PSUs and schools within PSUs selected in 1971-72 and 1978-79 by age. Appendix C gives information about the number of students assessed.

TABLE 1. Number of PSUs and Schools Within PSUs Selected in 1971-72 and 1978-79

	1971-72 Assessment		1978-79 Assessment	
	No. of PSUs	No. of Schools	No. of PSUs	No. of Schools
Age 9	116	777	75	648
Age 13	116	890	75	650
Age 17	116	780	75	534

CHAPTER 5

DATA COLLECTION

Professional data collection staff from the Research Triangle Institute, Raleigh, North Carolina, administered the assessment booklets. This staff was used to minimize the burden on participating schools and to ensure, insofar as possible, uniform administrative conditions across the country (Final Report...In-School Field Operations..., 1979).

Participation in the National Assessment is voluntary. NAEP makes every effort to encourage the schools selected in the sample to participate in the assessment, and National Assessment and Research Triangle Institute staffs have obtained high rates of school cooperation, as shown in Table 2 (Final Report...Field Operations..., 1979, p. 39, Table 27). Student cooperation rates were also high. The effect of student nonresponse is discussed in Appendix C. Table 3 shows the actual number of students that responded to a particular exercise booklet at each age level in the 1978-79 assessment.

TABLE 2. School Cooperation Rates, 1978-79 Assessment

Age	Percent of Eligible Schools Participating in 1978-79 Assessment
9	90.4
13	90.9
17	92.9
Overall	91.3

In 1978-79, booklets were administered to groups of 10-25 students; each group responded to only one of the booklets for their age level. The groups varied in size depending on an estimate of the rate of nonresponse for a particular school. In 1971-72, the planned session sizes were fixed at 12 students at each age.

In each assessment, steps were taken to guarantee the anonymity of respondents. Students' names were listed with their booklet identification number so that scoring and processing personnel could go back to the school lists for data verification -- for instance, on background information -- if necessary. These lists did not leave the schools and were destroyed six months following the assessment in a school.

TABLE 3. Number of Students Responding to Each Item Booklet in 1978-79 Assessment, by Age

Age 9		Age 13		Age 17	
Booklet	Number Responding	Booklet	Number Responding	Booklet	Number Responding
1	2,532	1	2,755	1	2,730
2	2,553	2	2,801	2+	2,746
3	2,475	3	2,775	3	2,761
4	2,494	4	2,791	4+	2,772
5	2,479	5	2,785	5	2,684
6	2,522	6	2,748	6	2,739
7	2,531	7	2,736	7	2,642
8	2,524	8	2,779	8	2,656
9+	2,486	9	2,754	9	2,787
10+	2,483	10+	2,758	10	2,697
11+	2,526	11+	2,751	11+	2,628
		12+	2,720	12+	2,628
		13+	2,757	13+	2,698
				14+	2,654
Total	27,605		35,910		37,822

+There were no music exercises included in booklets 9, 10 and 11 at age 9, booklets 10, 11, 12 and 13 at age 13, and booklets 2, 4, 11, 12, 13 and 14 at age 17.

To provide information on respondents' backgrounds, school officials were asked to respond to a "principal's questionnaire." This questionnaire asked whether music is taught as a required course or as an elective course, what type of music courses are available to students and the number of students who participate in music courses. Principals of 9-year-olds' schools were asked to indicate if music is taught and if so, by whom; the number of minutes per week each student receives music instruction; if instrumental music instruction is available and if so, the number of minutes per week that instruction is available. Forms used to collect background information from students and school officials are presented in Appendix B.

The assessment administrator coded each student's birth date, sex, grade, racial/ethnic classification and identification number on his or her booklet. Administrators made a visual racial/ethnic identification at the time each student turned in his or her booklet. During the 1978-79 assessment, six different racial classifications were used: white, black, Spanish heri-

tage, American Indian or Alaskan native, Pacific Islander or Asian, and unclassified. If an administrator was unsure of a student's racial/ethnic group, he or she referred to the student's name or listened to the student talk to make the identification. Students were not verbally asked by the assessment administrator to give a racial identification for themselves; however, 17-year-old students were asked to provide this information in one of the background questions included in the exercise booklet.

Sample sizes of the two classifications American Indian or Alaskan native and Pacific Islander or Asian are too small to permit reporting for these groups. Results for the group classified as Spanish heritage cannot be reported for separate exercises but can be reported for aggregate results across a number of exercises. Change results between 1971-72 and 1978-79 are reported for whites, blacks, and Hispanos or Spanish surnamed.

Each age group was assessed at approximately the same time of the school year in each assessment. As noted previously, 13-year-olds were assessed in October-December, 9-year-olds in January-February and 17-year-olds in March-May.

Following data collection, assessment administrators sent completed booklets to the scoring contractor, Westinghouse DataScore Systems, Iowa City, Iowa. Booklets were quality-checked to verify that correct administrative procedures were being followed by the field staff and that all booklets were accounted for. Coded identification information was also checked for accuracy; inconsistencies that could not be reconciled were sent back to the assessment administrator to be checked against the list of students' names and identification numbers retained by the school for six months following the assessment.

In 1971-72, 17-year-olds who were not currently attending school and young adults aged 26-35 were included in the assessment. Out-of-school 17-year-olds and adults could each answer up to four booklets of assessment materials; they were paid $5.00 for each booklet that they completed. Unpaced audio tapes were used for these groups. Assessment of out-of-school 17-year-olds and adults is quite expensive, and in 1978-79 funds to assess these groups were not available. Since out-of-school 17-year-olds were not assessed in 1978-79, only results for 17-year-olds attending school in 1971-72 and 17-year-olds attending school in 1978-79 were used in calculating changes in 17-year-olds' performance.

CHAPTER 6

SCORING

Data collected by National Assessment must be converted into a form suitable for computer processing and analysis. The conversion was done by Westinghouse DataScore Systems, Iowa City, Iowa, the contractor responsible for printing the assessment exercise booklets, receiving, scoring and machine-processing the data.

While most of the exercises in the 1978-79 music assessment were multiple-choice, seven exercises were open-ended. Responses to multiple-choice exercises were read directly from the booklets by optical scanning machines. The scoring contractor employed a special staff to hand score the open-ended exercises. Scorers were responsible for categorizing responses, using the scoring guides for open-ended exercises that defined categories of acceptable and unacceptable responses. Scorers were also responsible for coding the category of responses for an exercise into ovals that could be read by optical scanning machines.

To ensure accurate measurement of changes in performance, scoring procedures for open-ended exercises and correct answer keys for multiple-choice exercises had to be the same in each assessment. The open-ended exercises assessed in 1971-72 were reassessed in 1978-79, and the same scoring guides were used in both assessments. However, because of the complexity of one exercise, the 1971-72 and 1978-79 responses to this exercise were rescored to ensure that the same procedures were used and thus avoid inconsistency across scorers in different years.

Scorers were trained to use the guides by scoring sample responses taken from arriving assessment data. National Assessment staff was involved in the training process. Scorers initially worked as a group and discussed the appropriate categorization of each example response. Next, scorers worked individually on another set of responses. Discrepancies were resolved and explained. Once the group felt comfortable using the guides, they categorized the actual data. Supervisory personnel checked all work done for the first few days of a scoring effort to ensure consistency.

To further ensure the quality and consistency of scoring for open-ended exercises, quality-control checks were con-

ducted during the scoring of these exercises. At weekly intervals, randomly selected responses were drawn from the total pool of responses for an item and read by randomly selected scorers. Approximately 15% of the responses were included in the quality-control check. Scores for the quality-control readings were recorded, and the responses selected for quality control were then put back into the total pool of responses to be scored during the regular course of scoring. Following scoring of all responses, the two scores for quality-control responses were compared. If discrepancies in scoring became apparent, scorers were retrained and, on some occasions, work was rescored.

Percentages of agreement between quality-control and regular scoring were computed for each open-ended exercise. These data are summarized in Table 4.

TABLE 4. Quality-Control Summaries for the 1978-79 Music Assessment Open-Ended Exercises, by Age

	Exercise #	% of Scorer Agreement	% of Total Responses Sampled
Age 9	203001	95.6	15
	203015	99.2	15
	401003	97.9	25
Age 13	203001	95.1	15
	203015	97.9	15
	401003	97.5	25
	403001	100.0	15
	404003	99.0	15
Age 17	203001	97.6	15
	203015	98.7	15
	401003	95.8	15
	403001	100.0	15
	404003	100.0	15
	404004	99.5	15

CHAPTER 7

DATA ANALYSIS

Measures of Achievement

The basic measure of achievement reported by National Assessment is the percentage of respondents answering a given item acceptably. This percentage is an estimate of the percentage of 9-, 13- or 17-year-olds who would respond acceptably to a given item if every 9-, 13- or 17-year-old in the country were assessed.

Percentages of acceptable responses are used because each item is designed as a separate measure of some aspect of an objective or subobjective. The purpose of National Assessment is to discover if more or fewer people are able to answer these items acceptably -- and thus meet the objectives -- over time.

In addition to providing results on individual items, National Assessment reports the average performance across groups of similar items -- for the learning area as a whole, for a particular objective or subobjective, and so forth. Such a result constitutes the mean, or arithmetic, average of the estimates of performance on the individual items of the group and is called the mean percentage correct.[1] The items included in the calculation of a mean percentage usually are located in several exercise booklets, and thus the mean percentage should not be construed as an average test score.

To present a general picture of differences in achievement,

---

[1]Twenty-two empirical distributions of change measures from the 1969-70 and 1972-73 science assessments were used to generate Monte Carlo simulations of sampling distributions for several measures of central location. In addition to the mean and median, other measures of central location considered in the simulation studies included the average of the extremes, two forms of biweighted estimates and three forms of weight-matching estimation described by John W. Tukey in the research paper "Some Considerations on Locators Apt for Some Squeezed-Tail (and Stretched Tail) Parents" (1975). In almost every case, the sampling stability of the mean change was as good as or better than that of the other measures studied.

National Assessment reports the gains and losses on a group of exercises in terms of the differences in the average percentages of acceptable responses.

Unless the items summarized in the mean percentages of acceptable responses are identical, however, the means of one age group should not be compared directly with the means of another, since their values reflect both the choice of exercises and the performance of the students. When only a few exercises are summarized by a mean, one should be especially cautious in interpreting results, since a small set of exercises might not adequately cover the wide range of potential behaviors included under a given objective or subobjective. The mean should be interpreted literally as the arithmetic average of the percentage of acceptable responses obtained from National Assessment samples on a specific set of exercises, not as an average test score.

In addition to providing national results, National Assessment reports on the achievement of various subpopulations of interest. Groups are defined by region of the country, sex, race/ethnicity, size and type of community lived in, level of parents' education and grade in school (see Appendix A for definitions of these groups).

In considering National Assessment's achievement measures, differences in performance among ages and between assessments are the most useful. By maintaining the same item or set of items in making these comparisons, we have a reasonable indicator of whether more or fewer people know or can do something judged important.

The percentage difference between the performance of any one reporting group and that of the entire age group (nation) is recorded as a positive number if the group achieved a higher percentage or average than the entire age group and as a negative number if a lower. For example, a group performance of +1.8% indicates that the percentage of responses for the group is 1.8 percentage points higher than the national percentage of responses for that age level.

Procedures for estimating percentages of responses to exercises are dependent on the sample design. Each response by an individual was weighted and multiplied by an adjustment factor for nonresponse.[2] An estimate of the percentages of a particular age group that would have responded to a particular exercise in a particular way if the entire age group were assessed was defined as the weighted number of that type of response divided by the weighted number of all the responses. A

---

[2]Appendix C discusses nonresponse in assessment samples.

similar ratio of weights was used to estimate percentages and averages for reporting groups or subpopulations of interest.[3]

Estimating Variability in Achievement Measures

National Assessment used a national probability sample at each age level to estimate the proportion of people who would complete an exercise in a particular way. The sample selected was one of many possible samples of the same size that could have been selected using the same sample design. Since an achievement measure computed from each of the possible samples would differ, the standard error of the statistic was used as a measure of the sampling variability among achievement measures from all possible samples. A standard error, based on one particular sample, serves to estimate that sampling variability.

In the interest of sampling and cost efficiencies, National Assessment uses a complex, stratified, multistage probability sample design. Typically, complex designs do not provide for unbiased or simple computation of sampling errors. A reasonably good approximation of standard error estimates of acceptable response percentages and averages was obtained by applying the jackknife procedure (Miller, 1964, pp. 1594-1705; Miller, 1968, pp. 567-82; Mosteller and Tukey, 1968) to first-stage sampling units within strata. Standard errors for achievement measures such as national percentages, group differences and means or mean differences for a particular assessment year were estimated directly, taking advantage of features of the jackknife procedure that are generic to all of these statistics.[4] Since samples for different assessments are independent, the standard errors of the differences in achievement measures between assessments can be estimated simply by the square root of the sum of the squared standard errors from each of the assessments.

The standard error provides an estimate of sampling reliability for the achievement measures used by National Assessment. It is comprised of sampling error and other random error associated with the assessment of a specific item or set of items. Random error

---

[3]Following the 1978-79 assessment, a weighting-class adjustment procedure was used to smooth estimated population proportions across the 10 assessments conducted between 1969-70 and 1978-79. A discussion of this procedure is included in Appendix E.

[4]See Appendix D for a more detailed description of National Assessment's computation of standard errors.

includes all possible nonsystematic error associated with administering specific exercises to specific students in specific situations. Random differences among scorers for open-ended items are also included in the standard errors.

National Assessment has adhered to a standard convention whereby differences between statistics are designated as statistically significant at the .05 level of significance. That is, differences in performance between assessment years or between a reporting group and the nation are highlighted with asterisks only if they are at least twice as large as their standard error. Differences this large would occur by chance in fewer than 5% of all possible replications of our sampling and data collection procedures for any particular reporting group or national estimates.

Controlling Nonrandom Errors

Systematic errors can be introduced at any stage of an assessment -- exercise development, preparation of exercise booklets, design or administrative procedures, field administration, scoring or analysis. These nonsampling, nonrandom errors rarely can be quantified, nor can the magnitude of the bias they introduce into the estimates be evaluated directly.

Systematic errors can be controlled in large part by employing uniform administrative and scoring procedures and requiring rigorous quality control in all phases of an assessment. If the systematic errors are the same from age to age or group to group, then the differences in percentages or mean percentages are measures with reduced bias because subtraction tends to cancel the effect of the systematic errors.

Similarly, the effect of systematic errors in different assessment years can be controlled by carefully replicating in the second assessment the procedures carried out in the first. Differences in achievement across assessment years will also be measures with reduced bias since subtraction will again tend to cancel systematic errors.

Although it is not possible for every condition or procedure to remain exactly the same between assessments conducted several years apart, National Assessment has made every effort to keep conditions as nearly the same as possible. Changes in procedures described in this report were judged to have a relatively minor impact.

CHAPTER 8

REPORTS ABOUT THE MUSIC ASSESSMENT

Each assessment generates a tremendous amount of data. To make these data as useful as possible to a variety of audiences, National Assessment provides several types of publications and services.

Reports

Objectives

A description of the 1978-79 music objectives and the procedures used in developing the objectives and items for this assessment is available in: Music Objectives, Second Assessment.

Exercise-Level and Summary Data

For those wishing to use specific National Assessment items, NAEP provides copies of released items, exercise documentation (including exercise timing, objective being measured, administrative mode and source information for music stimuli used) and scoring guides for open-ended exercises. Data for the released items are provided as addendum sheets to the released exercise set. This loose-leaf set of materials is: The Second Assessment of Music, 1978-79: Released Exercise Set.

A cassette tape of the music stimuli for the released exercises from the 1978-79 music assessment is also available.

Overview Report

In addition to computing exercise-level data, National Assessment provides means for selected sets of exercises within a learning area assessment. These means are useful in comparing performance between assessment years and between reporting groups. Means are computed for a number of reporting variables, including region, sex, race/ethnicity, size and type of community, grade and level of parental education, as well as additional variables on the kinds of music courses taken by 9-, 13- and 17-year-olds.

Besides publishing many of these results in tabular form, National Assessment prepares overview reports for the general public -- including parents, classroom teachers, school administrators and legislators -- that summarize trends and highlight assessment results.

The report, Music 1971-79:

Results From the Second Music Assessment, describes the 1978-79 status in music achievement and changes between 1971-72 and 1978-79.

Public-Use Data Tapes and User Services

For those who wish to perform their own analyses of National Assessment data, computer data tapes of respondent-level data for the 1978-79 music assessment will be made available as funding levels permit.

National Assessment provides some assistance to those wishing to use the assessment items or to replicate assessment methodology. Those interested in the public-use data tapes or in receiving assistance should contact the National Assessment office.

APPENDIX A

DEFINITIONS OF NATIONAL ASSESSMENT REPORTING GROUPS

In addition to reporting results for all 9-, 13- and 17-year-old students in the United States, National Assessment reports results for a number of population subgroups. Most of these subgroups are defined for both the 1971-72 and 1978-79 music assessments.

Definitions of the subgroups follow:

Region

The country has been divided into four regions: Northeast, Southeast, Central and West. States included in each region are shown on the following map.

Sex

Results are reported for males and females.

Race/Ethnicity

Results are presented for blacks, whites and Hispanos.

Level of Parental Education

Three categories of parental-education levels are defined by National Assessment, based on students' reports. These categories are: (1) those whose parents did not graduate from high school, (2) those who have at least one parent who graduated from high school and (3) those who have at least one parent who has had some post-high-school education.

Type of Community

Communities in this category are defined by an occupational profile of the area served by a school as well as by the size of the community in which the school is located. This category is the only reporting category that excludes a large number of respondents. About two-thirds do not fall into

the classifications listed below. Results for the remaining two-thirds are not reported since their performance was similar to that of the nation.

Advantaged-urban (high-metro) communities. Students in this group attend schools in or around cities having a population greater than 200,000 where a high proportion of the residents are in professional or managerial positions.

Disadvantaged-urban (low-metro) communities. Students in this group attend schools in or around cities having a population greater than 200,000 where a relatively high proportion of the residents are on welfare or are not regularly employed.

Rural communities. Students in this group attend schools in areas with a population under 10,000 where many of the residents are farmers or farm workers.

Size of Community

Big cities. Students in this group attend schools within the city limits of cities having a 1970 census population over 200,000.

Fringes around big cities. Students in this group attend schools within metropolitan areas (1970 U.S. Bureau of the Census urbanized areas) served by cities having a population greater than 200,000 but outside the city limits.

Medium cities. Students in this group attend schools in cities having a population between 25,000 and 200,000, not classified in the fringes-around-big-cities category.

Small places. Students in this group attend schools in communities having a population less than 25,000, not classified in the fringes-around-big-cities category.

Grade in School

Results are categorized for 9-year-olds in the 3rd or 4th grade, 13-year-olds in the 7th or 8th grade, and 17-year-olds in the 10th, 11th or 12th grade.

Modal Grade by Region

Results are categorized for 9-, 13- and 17-year-old respondents in grades 4, 8 and 11, respectively, who live in the Northeastern, Southeastern, Central or Western region of the country.

Modal Grade by Community Size

Results are categorized for 9-, 13- and 17-year-old respondents in grades 4, 8 and 11, respectively, who live in big cities, fringes around big cities, medium cities and small places.

Modal Grade by Sex

Results are categorized for 9-, 13- and 17-year-old males and females in grades 4, 8 and 11, respectively.

Music Background Questions

Nine-year-olds were asked if they were being taught music and what kinds of activities they were being exposed to in their music classes. Thirteen- and 17-year-olds were asked about their music course-taking experiences in school. All three age groups were asked about their experiences with music outside of school. These music background questions are found in Appendix B.

APPENDIX B

FORMS USED TO OBTAIN BACKGROUND INFORMATION

This appendix includes the forms used by National Assessment to collect background information from school officials and respondents for the 1978-79 assessment. Following is a listing and a brief description of the forms included.

p. 31 School Principal's Questionnaire -- filled out by school principals or other school officials for schools at each of the age levels discussed.

p. 33 Principal's Music Questionnaire: Age 9 -- given to the elementary school principals. Provides information about the elementary school's music program.

p. 35 Principal's Music Questionnaire: Age 13 -- given to the principal at intermediate grades. Provides information about the music program in the intermediate grades.

p. 37 Principal's Music Questionnaire: Age 17 -- given to the principal in the senior high school. Provides information about the senior high school's music program.

p. 39 Standard Background Information Form for 9-Year-Olds -- provides information about reading material in the home and level of parents' education.

p. 41 Standard Background Information Form for 13-Year-Olds -- provides information about reading material in the home, level of parents' education and place lived in at age 9.

p. 43 Standard Background Information Form for 17-Year-Olds -- provides information on homework, TV watching, racial identification,

possessions in the home and classroom activities, in addition to questions also asked of 9- and 13-year-olds.

p. 47 Background Information on Music Experiences for 9-Year-Olds -- provides data on in-school music activities of 9-year-olds.

p. 50 Background Information on Music Experiences for 13- and 17-Year-Olds -- provides data on the in-school music courses taken by 13- and 17-year-olds.

p. 51 Background Information on Music Experiences for 9-, 13- and 17-Year-Olds -- provides data on the out-of-school music activities of these three age groups.

BEST COPY AVAILABLE

School Principal's Questionnaire

This report is authorized by law (20 U.S.C. 1221 c-1). While you are not required to respond, your cooperation is needed to make the results of this survey comprehensive, accurate, and timely.

Primary Sampling Unit ☐☐ School Number ☐☐☐☐☐☐☐☐☐

Age Group(s) 9 13 17

PLEASE PRINT

Name of School _____

Address of School _____
(Street)

_____
(City) (State) (Zip Code)

Name of School Principal _____

Name and title of person completing the form if other than school principal

Name _____ Title _____

1. What is your best estimate of the current enrollment and the average daily attendance by grade of your school (1978-79 school year)? (Enter zeros for grades not served by your school.)

Grade	K	1	2	3	4	5	6	7	8	9	10	11	12
Enrollment													
Average Daily Attendance													

2. Approximately what percentage of the students attending your school live in each of the following areas?

____% A In a rural area (less than 2,500)

____% B In a town of 2,500 to 10,000

____% C In a town of 10,000 or more

____
100% (Items A-C should add to 100%)

BEST COPY AVAILABLE

3. Approximately what percentage of the students attending your school are children of

_____% A Professional or managerial personnel

_____% B Sales, clerical, technical or skilled workers

_____% C Factory or other blue collar workers

_____% D Farm workers

_____% E Persons not regularly employed

_____% F Persons on welfare

_____ (Items A-F should add to 100%)
100%

4. Approximately what percentage of the students attending your school are

_____% A American Indian or Alaskan Native

_____% B Asian or Pacific Islander

_____% C Hispanic, regardless of race (Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin)

_____% D Black and not Hispanic

_____% E White and not Hispanic

_____ (Items A-E should add to 100%)
100%

5. Does your school qualify for ESEA Title I assistance?

_____ Yes - If Yes, approximately what number of students qualify for and what number of students are receiving ESEA Title I assistance?

_____ Approximate number of students qualifying for ESEA Title I assistance

_____ Approximate number of students receiving ESEA Title I assistance

_____ No

10. Is music taught in your school?

⬭ Yes ⬭ No (Stop. Do not answer the remaining questions.)

11. Who does the music instruction in your school?

⬭ The classroom teacher

⬭ A member of a teaching team having a background in music but not certified in music

⬭ A certified music teacher

⬭ A combination of classroom teacher and music specialist

⬭ Other

12. Where is music taught in your school?

⬭ In a regular classroom

⬭ In a special music room

⬭ Other

13. On the average, how much music instruction time each week do students receive?

⬭ Less than 30 minutes per week

⬭ 30 to 59 minutes per week

⬭ 60 to 89 minutes per week

⬭ 90 to 119 minutes per week

⬭ 120 minutes and over per week

14. Does your school offer instrumental music lessons for students?

⬭ Yes ⬭ No (Go to Question 16)

15. On the average, how much instruction time is given each week for instrumental music lessons for students?

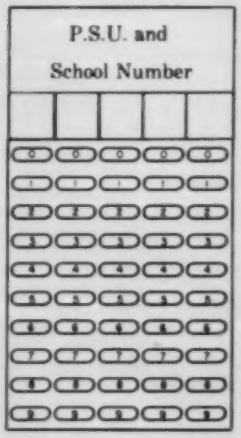
- ⬭ Less than 30 minutes per week
- ⬭ 30 to 59 minutes per week
- ⬭ 60 to 89 minutes per week
- ⬭ 90 to 119 minutes per week
- ⬭ 120 minutes and over per week

16. Do nine-year-old students in your school receive music instruction?

⬭ Yes ⬭ No (Stop)

17. On the average, how much time each week is given to music instruction of nine-year-old students?

- ⬭ Less than 30 minutes per week
- ⬭ 30 to 59 minutes per week
- ⬭ 60 to 89 minutes per week
- ⬭ 90 to 119 minutes per week
- ⬭ 120 minutes and over per week



THANK YOU FOR YOUR COOPERATION

School Name: _____

14. Is music instruction required for students in each of the following grades in your school?

	No such grade	Yes	No
A. Grade 6	⬭	⬭	⬭
B. Grade 7	⬭	⬭	⬭
C. Grade 8	⬭	⬭	⬭
D. Grade 9	⬭	⬭	⬭
E. Ungraded	⬭	⬭	⬭

(If the answers to ALL parts for Question 14 are NO, go to Question 16 on page 5.)

15. If music instruction is required for these grades, for what length of time do students receive music instruction?

	No such grade	Not required for this grade	1/4 school year or less	1/2 school year	3/4 school year	Full school year
A. Grade 6	⬭	⬭	⬭	⬭	⬭	⬭
B. Grade 7	⬭	⬭	⬭	⬭	⬭	⬭
C. Grade 8	⬭	⬭	⬭	⬭	⬭	⬭
D. Grade 9	⬭	⬭	⬭	⬭	⬭	⬭
E. Ungraded	⬭	⬭	⬭	⬭	⬭	⬭

16. Are elective music courses available for students in each of the following grades?

	No such grade	Yes	No
A. Grade 6	⬭	⬭	⬭
B. Grade 7	⬭	⬭	⬭
C. Grade 8	⬭	⬭	⬭
D. Grade 9	⬭	⬭	⬭
E. Ungraded	⬭	⬭	⬭

(If the answers to ALL parts for Question 16 are NO, go to Question 18 on page 6.)

17. If music is an elective, what percentage of the students in the following grades elect to take music courses each year?

	No such grade	Not elective for this grade	0-24%	25-49%	50-74%	75-89%	90-100%
A. Grade 6	⬭	⬭	⬭	⬭	⬭	⬭	⬭
B. Grade 7	⬭	⬭	⬭	⬭	⬭	⬭	⬭
C. Grade 8	⬭	⬭	⬭	⬭	⬭	⬭	⬭
D. Grade 9	⬭	⬭	⬭	⬭	⬭	⬭	⬭
E. Ungraded	⬭	⬭	⬭	⬭	⬭	⬭	⬭

7. Is music instruction, including music classes, group or private instruction, and performing groups, available in your school for the following grades?

		No such Grade	Yes	No
A.	Grade 9	⬭	⬭	⬭
B.	Grade 10	⬭	⬭	⬭
C.	Grade 11	⬭	⬭	⬭
D.	Grade 12	⬭	⬭	⬭
E.	Ungraded	⬭	⬭	⬭

8. Does your school offer the following types of music instruction at some time during Grades 9, 10, 11, or 12?

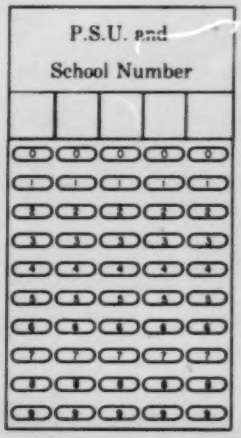
		Yes	No
A.	General Music class or Music Appreciation	⬭	⬭
B.	Music Literature, Music History or Introduction to Music	⬭	⬭
C.	Music Theory class or Music Composition	⬭	⬭
D.	Choir, Chorus, Glee Club or other vocal music group	⬭	⬭
E.	Vocal lessons (class or private)	⬭	⬭
F.	Band or Instrumental Music	⬭	⬭
G.	Stage or Jazz Band	⬭	⬭
H.	Jazz class or Improvisation class	⬭	⬭
I.	Orchestra	⬭	⬭
J.	Instrumental lessons (class or private)	⬭	⬭

9. About what percent of the 17-year-old students are currently enrolled in music classes (not including band, orchestra or vocal groups)?

- ⬭ 0 to 24%
- ⬭ 25 to 49%
- ⬭ 50 to 74%
- ⬭ 75 to 89%
- ⬭ 90 to 100%

10. About what percentage of the 17-year-old students are currently enrolled in performing groups, such as band, orchestra, or vocal music groups (choir, chorus, or glee clubs)?

- ⬭ 0 to 24%
- ⬭ 25 to 49%
- ⬭ 50 to 74%
- ⬭ 75 to 89%
- ⬭ 90 to 100%



THANK YOU FOR YOUR COOPERATION

School Name: _____

1. Does your family get a newspaper regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

2. Does your family get any magazines regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

3. Are there more than 25 books in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

4. Is there an encyclopedia in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

5. How much school did your father complete?
   (FILL IN THE ONE OVAL which best shows how much school your father completed.)

   ⬭ Did not complete the 8th grade

   ⬭ Completed the 8th grade, but did not go to high school

   ⬭ Went to high school, but did not graduate from high school

   ⬭ Graduated from high school

   ⬭ Some education after graduation from high school

   ⬭ I don't know.

6. Did your father graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

7. How much school did your mother complete?
(FILL IN THE ONE OVAL which best shows how much school your mother completed.)

- ⬭ Did not complete the 8th grade
- ⬭ Completed the 8th grade, but did not go to high school
- ⬭ Went to high school, but did not graduate from high school
- ⬭ Graduated from high school
- ⬭ Some education after graduation from high school
- ⬭ I don't know.

8. Did your mother graduate from a college or university?

⬭ Yes ⬭ No ⬭ I don't know.

1. Does your family get a newspaper regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

2. Does your family get any magazines regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

3. Are there more than 25 books in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

4. Is there an encyclopedia in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

5. How much school did your father complete?
   (FILL IN THE ONE OVAL which best shows how much school your father completed.)

   ⬭ Did not complete the 8th grade
   ⬭ Completed the 8th grade, but did not go to high school
   ⬭ Went to high school, but did not graduate from high school
   ⬭ Graduated from high school
   ⬭ Some education after graduation from high school

   ⬭ I don't know.

6. Did your father graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

7. How much school did your mother complete?
   (FILL IN THE ONE OVAL which best shows how much school your mother completed.)

   ⬭ Did not complete the 8th grade
   ⬭ Completed the 8th grade, but did not go to high school
   ⬭ Went to high school, but did not graduate from high school
   ⬭ Graduated from high school
   ⬭ Some education after graduation from high school

   ⬭ I don't know.

8. Did your mother graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

9. **Where did you live on your ninth birthday?**

- ⬭ **In the United States (Please specify the state or territory.)**

  _____

- ⬭ **Outside the United States (Please specify the country.)**

  _____

- ⬭ **I don't know.**

1. Which of the following does your family have at home? (Fill in one oval on each line.)

		Have	Do not have
A.	Newspaper received regularly	⬭	⬭
B.	Magazines received regularly	⬭	⬭
C.	More than 25 books	⬭	⬭
D.	Encyclopedia	⬭	⬭
E.	Dictionary	⬭	⬭
F.	Record player	⬭	⬭
G.	Tape recorder or cassette player	⬭	⬭
H.	Typewriter	⬭	⬭
I.	Vacuum cleaner	⬭	⬭
J.	Electric dishwasher	⬭	⬭
K.	Two or more cars or trucks that run	⬭	⬭

2. How much time did you spend on homework yesterday?

⬭ No homework was assigned
⬭ I had homework but didn't do it
⬭ Less than one hour
⬭ Between 1 and 2 hours
⬭ More than 2 hours

3. How many different schools have you attended since you started the first grade?

⬭ 1 to 3 schools
⬭ 4 to 6 schools
⬭ 7 to 9 schools
⬭ 10 or more schools

4. How long have you lived in the community in which you now live?

- ⬭ All my life
- ⬭ 10 or more years but not all my life
- ⬭ 5 to 9 years
- ⬭ 2 to 4 years
- ⬭ 1 year
- ⬭ Less than 1 year

5. How much television did you watch yesterday?

⬭ None	⬭ 2 hours	⬭ 5 hours
⬭ 1 hour or less	⬭ 3 hours	⬭ 6 hours or more
⬭ 1 hour	⬭ 4 hours	

6. Is English the language spoken most often in your home?

⬭ Yes ⬭ No

7. Is a language other than English spoken in your home?

⬭ Often ⬭ Sometimes ⬭ Never

8. How many brothers or sisters do you have who are <u>older</u> than you?

None	1	2	3	4	5	6 or more
⬭	⬭	⬭	⬭	⬭	⬭	⬭

9. How many brothers or sisters do you have who are <u>younger</u> than you?

None	1	2	3	4	5	6 or more
⬭	⬭	⬭	⬭	⬭	⬭	⬭

10. A. What is your racial background?

- ⬭ American Indian or Alaskan Native
- ⬭ Asian or Pacific Islander
- ⬭ Black
- ⬭ White
- ⬭ Other (Please specify) _____

B. Is your ethnic heritage Hispanic (such as Mexican, Puerto Rican, Cuban, Central or South American or other Spanish Culture or origin)?

⬭ Yes ⬭ No

11. How often has each of the following been used in the courses you are taking this year? (Fill in one oval on each line.)

		Never	Seldom	Fairly Often	Frequently
A.	Listening to the teacher's lecture	⬭	⬭	⬭	⬭
B.	Participating in student-centered discussions	⬭	⬭	⬭	⬭
C.	Working on a project or in a laboratory	⬭	⬭	⬭	⬭
D.	Writing essays, themes, poetry, stories	⬭	⬭	⬭	⬭
E.	Going on field trips	⬭	⬭	⬭	⬭
F.	Having individualized instruction (small groups or one-to-one with a teacher)	⬭	⬭	⬭	⬭
G.	Using teaching machines or computer-assisted instruction	⬭	⬭	⬭	⬭
H.	Watching television lectures	⬭	⬭	⬭	⬭
I.	Studying from textbooks	⬭	⬭	⬭	⬭
J.	Library or media center assignments	⬭	⬭	⬭	⬭

A. How much school did your father complete?
(FILL IN THE ONE OVAL which best shows how much school your father completed.)

⬭ Did not complete the 8th grade
⬭ Completed the 8th grade, but did not go to high school
⬭ Went to high school, but did not graduate from high school
⬭ Graduated from high school
⬭ Some education after graduation from high school

⬭ I don't know.

B. Did your father graduate from a college or university?

⬭ Yes ⬭ No ⬭ I don't know.

C. How much school did your mother complete?
(FILL IN THE ONE OVAL which best shows how much school your mother completed.)

- ⬭ Did not complete the 8th grade
- ⬭ Completed the 8th grade, but did not go to high school
- ⬭ Went to high school, but did not graduate from high school
- ⬭ Graduated from high school
- ⬭ Some education after graduation from high school
- ⬭ I don't know.

D. Did your mother graduate from a college or university?

⬭ Yes ⬭ No ⬭ I don't know.

E. Where did you live on your ninth birthday?

- ⬭ In the United States (Please specify the state or territory.)

  _____

- ⬭ Outside of the United States (Please specify the country.)

  _____

- ⬭ I don't know.

F. Where did you live on your thirteenth birthday?

- ⬭ In the United States (Please specify the state or territory.)

  _____

- ⬭ Outside of the United States (Please specify the country.)

  _____

- ⬭ I don't know.

Background Information on Music Experiences for 9-Year-Olds

A. Are you being taught music in school <u>this year</u>?

- ⬭ Yes
- ⬭ No
- ⬭ I don't know.

B. Were you taught music in school <u>last year</u>?

- ⬭ Yes
- ⬭ No
- ⬭ I don't know.

C. Do you ever do each of the following activities in your school music class? Fill in one oval in each box on this and the next page.

In your school music class, do you

1.	listen to music?	Yes ⬭	No ⬭	I don't know. ⬭
2.	sing just for fun?	Yes ⬭	No ⬭	I don't know. ⬭
3.	sing in a special music group such as a choir, chorus or glee club?	Yes ⬭	No ⬭	I don't know. ⬭
4.	play a musical instrument just for fun?	Yes ⬭	No ⬭	I don't know. ⬭
5.	play a musical instrument in a special music group such as a band or orchestra?	Yes ⬭	No ⬭	I don't know. ⬭
6.	read about music or musicians?	Yes ⬭	No ⬭	I don't know. ⬭
7.	make up your own music?	Yes ⬭	No ⬭	I don't know. ⬭

D. Which one of the following things would you rather do if you had one free period a day in school? Choose only one.

- ⬭ Play a musical instrument
- ⬭ Draw or paint
- ⬭ Write a story
- ⬭ Sing in a musical group
- ⬭ Learn a foreign language
- ⬭ Listen to music
- ⬭ None of these
- ⬭ I don't know.

Background Information on Music Experiences for 13- and 17-Year-Olds

Listed below are several kinds of music activities. Indicate how many years you have participated in each activity in school. Be sure to fill in one oval in each box.

Including elementary and junior high school, how many years have you taken or participated in

A. General Music Class or Music Appreciation?

Have Never Taken	Less Than 1 Year	1 to 2 Years	3 to 4 Years	5 or More Years
⬭	⬭	⬭	⬭	⬭

B. Choir, Chorus or Glee Club?

Have Never Taken	Less Than 1 Year	1 to 2 Years	3 to 4 Years	5 or More Years
⬭	⬭	⬭	⬭	⬭

C. Band or Instrumental Music?

Have Never Taken	Less Than 1 Year	1 to 2 Years	3 to 4 Years	5 or More Years
⬭	⬭	⬭	⬭	⬭

D. Orchestra?

Have Never Taken	Less Than 1 Year	1 to 2 Years	3 to 4 Years	5 or More Years
⬭	⬭	⬭	⬭	⬭

E. Introduction to Music, Music Literature or Music History?

Have Never Taken	Less Than 1 Year	1 to 2 Years	3 to 4 Years	5 or More Years
⬭	⬭	⬭	⬭	⬭

F. Music Theory Class or Music Composition?

Have Never Taken	Less Than 1 Year	1 to 2 Years	3 to 4 Years	5 or More Years
⬭	⬭	⬭	⬭	⬭

Background Information on Music Experiences for 9-, 13- and 17-Year-Olds

Which of the following activities do you do outside of school? Fill in one oval in each box.

Outside of school, do you

		Yes	No	I don't know.
A.	listen to music?	⬭	⬭	⬭
B.	sing just for fun?	⬭	⬭	⬭
C.	sing with friends just for fun?	⬭	⬭	⬭
D.	sing in a church or community music group?	⬭	⬭	⬭
E.	play a musical instrument by yourself just for fun?	⬭	⬭	⬭
F.	play a musical instrument with friends just for fun?	⬭	⬭	⬭
G.	play a musical instrument in a church or community music group?	⬭	⬭	⬭
H.	take music lessons?	⬭	⬭	⬭
I.	make up your own music?	⬭	⬭	⬭

APPENDIX C

RESPONSE RATES FOR ASSESSMENT SAMPLES

Table C-1 shows the response rates for students assessed in 1971-72 and 1978-79. In the 1971-72 assessment, for each of the three age groups, 12 students and 4 alternates were selected for each assessment session. If all 12 students appeared for the session, then the alternates were dismissed. Otherwise, enough alternates were selected to bring the size of the group to as many students as possible up to a total of 12. If the group assessed numbered between 8 and 12 students, then the administration was considered complete. If the final total was not at least a quorum of 8, a second and sometimes a third make-up session was held. The percentages in Table C-1 are based on the numbers of students assessed from the original groups of 12 selected and do not reflect the use of alternates.

TABLE C-1. Number of Students Assessed and Percent of Sample Covered, by Age and Assessment Year

Year	Age	Type of Administration	Number of Packages	Total Number of Students Assessed	Average Number Assessed Per Package	Average Sample Coverage in Percent
1971-72	9	G	4	10,824	2,706	89.2
		I	3	6,953	2,318	90.0
	13	G	7	18,669	2,667	87.0
		I	3	6,870	2,290	88.5
	17	G	9	21,233	2,359	78.2
		I	3	6,565	2,188	84.1
1978-79	9	G	11	27,605	2,510	87.8
	13	G	13	35,910	2,762	84.9
	17	G	14	37,822	2,702	77.8

+In the 1971-72 assessment, some booklets were administered to individual (I) students using an interview mode. The other booklets were administered to groups (G) of students.

For the 1978-79 assessment, slightly different procedures were used. The number of students selected for each administrative session varied from 16-25 students depending on previous response rates obtained from schools in similar communities. No alternates were selected. The quorum size needed to consider an administrative session complete varied according to the number of students selected. Since nonresponse rates have always been relatively small for ages 9 and 13, the make-up or follow-up procedures used in 1978-79 for these ages were similar to the ones used for the first music assessment. If a quorum was not obtained at the first administrative session, a second and sometimes a third make-up session was held. At age 17, in the 1978-79 assessment, follow-up procedures were conducted on a school, rather than a session, basis. If a school had an overall response rate of less than 75%, then all nonrespondents in the school were contacted for one or two follow-up sessions. These follow-up procedures for 17-year-olds provided sample coverage similar to that obtained at ages 9 and 13.

Since response rates at age 17 have always been somewhat lower than at the other two ages, the Research Triangle Institute (RTI), Raleigh, North Carolina, was asked to conduct a special study of nonrespondents during the 1972-73 assessment of science and mathematics. The results (Kalsbeek et al., 1975; Rogers et al., 1977) indicated that about 80% of the total nonrespondent group did not appear at the assessment sessions because of conflicting school activities or illness. The remaining 20% did not seem to be available. They attended school infrequently, if at all (for practical purposes, they had dropped out), or they had moved out of the school attendance area. In either case, these students probably should not have been listed in the in-school population of eligibles.

Tables published in previous National Assessment reports showing response rates for age 17 generally contain percentages adjusted to account for those 17-year-olds listed, but not attending school. But, since National Assessment has not had the resources to replicate the RTI study in recent assessments, the 20% figure used as a basis for adjusting these percentages may be outdated and thus, the percentages given in Table C-1 have not been adjusted. It seems likely that despite efforts to update the lists of eligibles, these lists still contain some percentage of students who have in effect left the schools. Thus, the percentages listed for age 17 are probably underestimates of the actual response rates for 17-year-olds attending school.

APPENDIX D

COMPUTATION OF MEASURES OF ACHIEVEMENT, CHANGES IN ACHIEVEMENT AND STANDARD ERRORS

Several measures of achievement that National Assessment uses in its reports are described in Chapter 7 of this document. The sample design, as described in Chapter 4, is a complex, deeply stratified, multistage probability sample design. Measures of achievement are obtained by weighting individual responses appropriately. Reasonably good approximation of standard error estimates of these achievement measures can be obtained by applying the jackknife procedure to first-stage sampling units within strata, using the method of successive differences and accumulating across strata.

In this section, the measures of achievement are first defined in algebraic form, followed by a description of the jackknife method that National Assessment uses to estimate the standard errors of achievement measures.

Measures of Achievement

Based on the sample design, a weight is assigned to every individual who responds to an exercise administered in an assessment. The weight is the reciprocal of the probability of selecting a particular individual to take a particular exercise with adjustment for nonresponse. Since the probabilities of selection are based on an estimated number of people in the target age population, the weight for an individual estimates the number of similar people that individual represents in the age population. As explained in Appendix E, the weights are adjusted to reflect information from previous assessments on population distributions.

A sum of the weights for all individuals at an age level responding to an exercise is an estimate of the total number of people in that age population. A sum of weights for all individuals at an age responding correctly to an exercise is an estimate of the number of people in the age population who would be able to respond correctly if the entire population were assessed. These concepts also apply to any reporting group (e.g., defined by region, sex, and so on) and category of response (e.g., correct, incorrect and "I don't know").

Let W^{e}_{ihk} = sum of weights for respondents to exercise e who are in reporting subgroup i and who are in the kth replicate of the hth sampling stratum, and

C^{ej}_{ihk} = sum of weights for respondents to exercise e who are in reporting subgroup i, who are in the kth replicate of the hth sampling

stratum and who selected response category j (e.g., correct foil) for the exercise.

Note that $W^e_{ihk} = \sum_j C^{ej}_{ihk}$.

Then summing k over the n_h sample replicates in the stratum h, and summing over the H sampling strata,

$$W^e_{i++} = \sum_{h=1}^{H} \sum_{k=1}^{n_h} W^e_{ihk}$$

estimates the number of eligibles in the population who are in subgroup i.

Similarly, $C^{ej}_{i++} = \sum_{h=1}^{H} \sum_{k=1}^{n_h} C^{ej}_{ihk}$ estimates the number of eligibles in the population who are in subgroup i and who would select response category j for exercise e.

An estimate of the proportion of the eligibles in the age population in group i who would select response category j on exercise e is:

(1) $P^{ej}_i = C^{ej}_{i++}/W^e_{i++}$.

In the special case where the proportion of all age-eligibles who would select response category j on exercise e is estimated, the index A (for ALL) will be used in place of i as follows:

(2) $P^{ej}_A = C^{ej}_{A++}/W^e_{A++}$.

In National Assessment reports, the proportion in (1) multiplied by 100 is called the group percentage, and the proportion in (2) multiplied by 100 is called the national percentage. The difference between the proportion in subgroup i who would select category j on exercise e and the proportion in the nation is denoted by:

(3) $\Delta P^{ej}_i = P^{ej}_i - P^{ej}_A$.

National Assessment also reports the arithmetic mean of the percentage of correct responses over sets of exercises corresponding to the measures in (1), (2) and (3). These means are taken over the set of all exercises or a subset of exercises classified by a reporting topic or content objective. The mean percentages of correct responses taken over m exercises in some set of exercises corresponding to measures (1), (2) and (3) are, respectively:

(4) $\bar{P}_i = \frac{1}{m} \sum_e C^e_{i++}/W^e_{i++}$,

(5) $\bar{P}_A = \frac{1}{m} \sum_e C^e_{A++}/W^e_{i++}$ and

(6) $\overline{\Delta P}_i + \bar{P}_i - \bar{P}_A$.

Note that the response category subscript j has been suppressed since the means are understood to be taken over the correct response category for each exercise.

Each of these six achievement measures is computed and routinely used in reports describing achievement data for any assessment. The simple difference in these measures between two assessments of the same exercise (or sets of exercises) provides six measures of change in achievement that are routinely used in National Assessment's change reports. The next section describes how standard errors are estimated for the 12 statistics

used in NAEP reports.

Computation of Standard Errors

In order to obtain an approximate measure of the sampling variability in the statistics (1) through (6), a jackknife replication procedure for estimating the sampling variance of nonlinear statistics from complex, multistage samples was tailored to National Assessment's sample design. Miller (1968, 1974) and Mosteller and Tukey (1977) provide information about the jackknife technique, while Folsom (1977) describes how the procedure is used in estimating standard errors for National Assessment's sample design.

To demonstrate the computational aspects of this technique, consider estimating the variance of the statistic in (1) -- the proportion of age-eligibles in subgroup i who would select response category j on exercise e.

This statistic is based on the data from all the n_h replicates in the H strata. Let P^{ej}_{i-hk} be defined as a replication estimate of P^{ej}_i and constructed from all the replicates excluding the data from replicate k in stratum h. These replication estimates are computed as if the excluded replicate had not responded, and a reasonable nonresponse adjustment is used to replace the data in replicate hk in estimating P^{ej}_i. Several choices for replacing the data in replicate hk are available. In order to obtain a convenient and computationally efficient algorithm for approximating standard errors, National Assessment replaces C^{ej}_{ihk} and W^{e}_{ihk} from the hkth replicate with corresponding sums from another paired replicate in the same stratum. The replicate estimate is then computed. The replicate estimates to be used in the calculations are determined by arranging all the replicates in each stratum into successive pairs. That is, replicate 1 is paired with replicate 2, replicate 2 with replicate 3, 3 with 4,...(n_h-1) with n_h and replicate n_h with replicate 1.

The contribution to the variance of P^{ej}_i by each pair of replicates is the change in the value of the statistic incurred by replacing the data from each replicate in the pair with the data from the other replicate in the pair and recomputing P^{ej}_i in the usual way. This produces two replicate estimates. Squaring the difference between these replicate estimates and then dividing by eight measures the contribution of this pair of replicates to the total variance. The sum of these contributions over all n_h successive pairs in the stratum is the contribution by stratum h to the total variance. The square root of the sum of the H stratum contributions is the estimate of the standard error of P^{ej}_i.

Algebraically, the two replicate estimates for the pair k, $k+1$ (where $k=1,\ldots n_h$ and $n_h+1=1$) are:

$$(7)\quad P^{ej}_{i\text{-}hk} = \frac{C^{ej}_{i++} - C^{ej}_{ihk} + C^{ej}_{ih(k+1)}}{W^{ej}_{i++} - W^{ej}_{ihk} + W^{ej}_{ih(k+1)}}$$

and

$$(8)\quad P^{ej}_{i\text{-}h(k+1)} = \frac{C^{ej}_{i++} - C^{ej}_{ih(k+1)} + C^{ej}_{ihk}}{W^{ej}_{i++} - W^{ej}_{ih(k+1)} + W^{ej}_{ihk}}.$$

The contribution to the total variance from stratum h is:

$$(9)\quad \text{var}\,(P^{ej}_{ih}) = \frac{1}{8}\sum_{k}^{n_h}\left(P^{ej}_{i\text{-}hk} - P^{ej}_{i\text{-}h(k+1)}\right)^2.$$

And finally, an estimate of the standard error of P^{ej}_{i} is:

$$(10)\quad SE\,(P^{ej}_{i}) = \left(\sum_{h}^{H} \text{var}\; P^{ej}_{ih}\right)^{1/2}.$$

Multiplying P^{ej}_{i} by 100 yields the percentage of response to category j. Multiplying $SE(P^{ej}_{i})$ by 100 yields the corresponding estimated standard error of the percentage.

In general, the jackknifed standard errors of the proportion estimates will be larger than the simple random sampling formula $(PQ/N)^{1/2}$, where $P=P^{ej}_{i}$, $Q=1-P$ and N is the number of sampled respondents in subgroup i who took the exercise. The larger size of $SE(P^{ej}_{i})$ reflects mainly the loss of precision due to cluster-sampling of schools and students. The standard errors for the achievement measures (2) through (6) are computed through a series of steps analogous to those followed in computing $SE(P^{ej}_{i})$.

The standard errors for the differences between two assessments for any of the achievement measures (1) through (6) are computed as the square root of the sum of the squared standard errors from each of the separate assessments.

The size of the standard errors depends largely not only on the number of replicates and schools included in the sample, but also on the number of respondents in each of the reporting groups. Table D-1 shows the average number of students responding to an exercise booklet for each of the reporting groups for each age for each assessment year. Table D-2 shows National Assessment's current estimates of the proportions of students in each reporting group at each age.

TABLE D-1. Average Number of Respondents in Reporting Groups Taking an Item Booklet, by Age and Assessment Year+

	Age 9		Age 13		Age 17#	
	1971-72	1978-79	1971-72	1978-79	1971-72	1978-79
Nation	2,538	2,510	2,584	2,762	2,317	2,702
Region						
Northeast	627	580	629	675	563	642
Southeast	631	625	658	657	595	683
Central	653	665	658	752	589	725
West	626	639	638	681	569	649
Sex						
Male	1,255	1,255	1,297	1,370	1,126	1,312
Female	1,283	1,255	1,286	1,395	1,191	1,386
Race/ethnicity						
White	1,977	1,849	2,069	2,053	1,890	2,134
Black	387	484	340	507	288	392
Other++	174	177	175	202	139	176
Parental education						
Not graduated high school	282	191	436	326	468	386
Graduated high school	607	589	789	895	715	895
Post high school	840	880	998	1,200	1,006	1,334
Unknown++	809	850	361	341	128	87
Type of community						
Rural	252	252	258	271	233	246
Disadvantaged urban	253	250	261	281	232	306
Advantaged urban	253	253	256	277	235	265
Other ++	1,780	1,755	,809	1,933	1,617	1,885
Size of community						
Big cities	484	723	510	775	435	718
Fringes around big cities	526	448	561	588	507	571
Medium cities	343	237	348	287	303	287
Small places	1,184	1,097	1,164	1,115	1,071	1,123
Grade						
3, 7, 10	644	623	667	683	280	360
4, 8, 11	1,807	1,834	1,799	2,031	1,681	2,014
12	—	—	—	—	255	285
Other++	87	53	118	48	101	43

+Data may not total due to rounding error.
#Seventeen-year-olds enrolled in school.
++Data are not reported for these groups.

TABLE D-2. Estimated Current Population Proportions of National Assessment Reporting Groups for In-School Students

Reporting Groups	Age 9	Age 13	Age 17
Sex			
Male	.50	.50	.48
Female	.50	.50	.52
Race/ethnicity			
White	.79	.80	.83
Black	.14	.13	.12
Other	.07	.07	.05
Region			
Northeast	.25	.25	.25
Southeast	.22	.23	.20
Central	.27	.27	.29
West	.26	.25	.26
Parental education			
Not graduated high school	.09	.13	.15
Graduated high school	.24	.32	.32
Post high school	.33	.42	.48
Unknown	.34	.13	.05
Type of community			
Rural	.08	.10	.08
Disadvantaged urban	.07	.07	.09
Advantaged urban	.11	.11	.11
Other	.74	.72	.72
Size of community			
Big cities	.20	.21	.19
Fringes around big cities	.22	.22	.26
Medium cities	.12	.11	.11
Small places	.46	.46	.44
Grade in school			
<3, <7, <10	<.01	.02	.02
3, 7, 10	.23	.25	.13
4, 8, 11	.75	.72	.75
>4, >8, 12	<.01	<.01	.10
Other	<.01	<.01	<.01

APPENDIX E

ADJUSTMENT OF RESPONDENT WEIGHTS BY SMOOTHING TO REDUCE RANDOM VARIABILITY OF ESTIMATED POPULATION PROPORTIONS

Background

As noted elsewhere, a weight is assigned to every individual who responds to an exercise administered in an assessment. The weight is the reciprocal of the probability of selection of the individual with adjustment for nonresponse, and the probabilities of selection are based on the estimated number of people in the target age populations. Therefore, the weight for an individual estimates the number of people that the individual represents in the age population. The sum of the weights of all individuals at an age level who responded to an exercise is an estimate of the total number of people in that age population in the year that the exercise was assessed. Similarly, the sum of weights for all individuals who took the exercise and who also are members of some demographic category (such as blacks) gives an estimate of the number of people in the age population, for the year, who are also members of the category. The ratio of the two totals estimates the proportional representation of the demographic category in the age population for the given year.

Separate estimates of the proportional representation of the various demographic subgroups are provided by each booklet administered to a particular age group in a given year. Due to random sampling variability, the estimates of population proportions for a given year based on single booklets will vary. There is also random sampling variation in estimates of population proportions from year to year in addition to whatever trends in population proportions over time that might exist.

It is desirable to reduce the random variability of population proportions as much as possible since this variability has an effect on performance estimates. For example, the percentage of acceptable responses for an age group is a function of the relative proportions of high-performing and low-performing groups. If the relative proportions of these groups are very different in different assessments due to sampling variability, then a portion of the change in percentage of acceptable responses for an age group could be attributable to yearly sampling differences in the relative proportions of

high- and low-achieving groups.

In addition to reporting performance estimates for an age group as a whole, National Assessment also reports performance for various subpopulations, such as whites or blacks. Because variability of subgroups within these subpopulations (such as males and females within the white subpopulation) influences the performance estimates for the subpopulations, it is desirable that fluctuations of proportions of all subgroups of each subpopulation be reduced as much as possible.

For each age and year, each of the various booklets administered provides estimates of a given population proportion. Since these estimates are subject to booklet-to-booklet variability, a better estimate of the population proportion, which will have reduced variability, is obtained by combining the information from all booklets. However, these proportions vary from year to year due to random sampling variability or systematic differences in sampling procedures. An even better estimate of population proportions for any single year can be obtained by smoothing the proportions over several assessment years. The word "smoothing" is used here in the sense of fitting a smooth curve to a sequence of numbers by robust/resistant procedures (Tukey, 1977). Smoothing estimates of population proportions reduces a large portion of the sampling variability while preserving, as far as possible, actual trends occurring in the age population.

After the population proportions have been smoothed, adjusted weights are derived for the assessed individuals so that the population proportions computed using the adjusted weights are equal to the smoothed proportions. The adjusted weights are then used for all analyses.

Smoothing Procedures Used by National Assessment

The most direct way to smooth proportions is first to classify people into mutually exclusive multiway cells on the basis of their membership in categories of various important variables and then to smooth the proportions within each of the resulting multiway cells across years. Unfortunately, this procedure tends to produce a large number of cells with few people and, consequently, quite unstable estimates of smoothed proportions.

To circumvent this difficulty, National Assessment has utilized various smoothing procedures. These procedures, which are all basically weighting-class adjustments applied independently to each age, are designed to control, to varying degrees, fluctuations in certain key subgroups while avoiding, as much as possible, instabilities due to small cells.

The procedure used in 1978-79 has the following characteristics:

1. It produces a single adjusted weight for each individual.

2. It affords good control on the distribution of proportions of certain key variables.

3. It tends to produce stability of performance estimates.

4. It is relatively easy to implement.

Even though adjusted weights using this procedure differ slightly from the corresponding adjusted weights from the other procedures that have previously been employed, National Assessment intends to use weights obtained using the 1978-79 procedure for all future analyses of data assessed in earlier years. This is simply because NAEP believes weights obtained through this procedure to be the best available.

The Current Smoothing Procedure

The first step in the 1978-79 smoothing procedure involved the partitioning of the population of age-class eligibles into the six smoothing cells given in Table E-1. The same cells were used for all ages.

Then, for each age and every year, the proportion of the population in each of the cells was estimated. For a given age and year, the proportion of the population in a particular cell was computed as the sum of weights of all respondents assessed in the given year who were of the specified age and who belonged in the cell, divided by the total of the weight of all respondents of the given age assessed in that year.

Each of the six cells was comprised of a sequence of estimated population proportions corresponding to the various years of assessment. Each such sequence of proportions was then smoothed by fitting robust/resistant lines. Using

---

TABLE E-1. Smoothing Cells Used for the 1978-79 Smoothing Procedure

Cell	Race	Region	Community Size
1	White	All	Big city + fringe
2	White	All	Medium city
3	White	All	Small places
4	Black	Southeast	All
5	Black	Not Southeast	All
6	Other	All	All

data from the U.S. Census and Current Population Survey, trends in enrollment by age and race and by age and region were obtained. The data from these surveys were adjusted to correspond with National Assessment definitions as much as possible. The resistant lines within the smoothing cells were constrained to satisfy the trend from the U.S. Census and Current Population Survey data.

The final step in the smoothing procedure was to adjust the respondents' weights to be consistent with the smoothed proportions. Since each respondent takes only one booklet, the weight adjustments were done independently for each booklet. For a given age, year and booklet, population proportions using the original weights were obtained for each of the smoothing cells. Then the weights of all respondents of a given cell were multiplied by the ratio of the smoothed cell proportion to the proportion using the original weights. This produced the adjusted weights that were used in all analyses.

Adjustment of Weights by Users

The smoothed population proportions for 9-, 13- and 17-year-olds (in-school only) are given in Tables E-2, E-3 and E-4, respectively. The columns

---

TABLE E-2. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 9-Year-Olds

	Cell					
	1	2	3	4	5	6
Race	White	White	White	Black	Black	Other
Region	All	All	All	SE+	Not SE	All
Size of Community	BC+FR#	MC++	SP##	All	All	All
Year						
1969-70	0.3293	0.1258	0.3546	0.0535	0.0745	0.0624
1970-71	0.3251	0.1224	0.3601	0.0547	0.0746	0.0631
1971-72	0.3210	0.1191	0.3656	0.0558	0.0747	0.0639
1972-73	0.3168	0.1158	0.3711	0.0570	0.0748	0.0646
1973-74	0.3126	0.1124	0.3766	0.0581	0.0748	0.0654
1974-75	0.3085	0.1091	0.3822	0.0592	0.0749	0.0661
1975-76	0.3043	0.1058	0.3877	0.0604	0.0750	0.0668
1976-77	0.3001	0.1024	0.3932	0.0615	0.0751	0.0676
1977-78	0.2959	0.0991	0.3987	0.0627	0.0752	0.0683
1978-79	0.2918	0.0958	0.4042	0.0638	0.0753	0.0691

+SE = Southeast.
#BC+FR = Big Cities + Fringes.
++MC = Medium Cities.
##SP = Small Places.

TABLE E-3. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 13-Year-Olds

	Cell					
	1	2	3	4	5	6
Race	White	White	White	Black	Black	Other
Region	All	All	All	SE+	Not SE	All
Size of Community	BC+FR#	MC++	SP##	All	All	All
Year						
1969-70	0.3200	0.1309	0.3703	0.0513	0.0673	0.0602
1970-71	0.3201	0.1256	0.3731	0.0521	0.0681	0.0610
1971-72	0.3202	0.1202	0.3760	0.0528	0.0690	0.0618
1972-73	0.3203	0.1149	0.3788	0.0536	0.0698	0.0626
1973-74	0.3203	0.1096	0.3816	0.0544	0.0714	0.0642
1974-75	0.3205	0.1043	0.3845	0.0552	0.0714	0.0642
1975-76	0.3206	0.0989	0.3873	0.0560	0.0722	0.0650
1976-77	0.3207	0.0936	0.3901	0.0567	0.0731	0.0658
1977-78	0.3207	0.0883	0.3929	0.0575	0.0739	0.0666
1978-79	0.3208	0.0830	0.3958	0.0583	0.0747	0.0674

+SE = Southeast.
#BC+FR = Big Cities + Fringes.
++MC = Medium Cities.
##SP = Small Places.

---

of each table represent the smoothing cells while the rows represent the assessment year. For example, the smoothed population proportion of 9-year-olds in smoothing cell 2 (whites in medium cities) for 1972-73 is .1158.

To adjust respondent weights to be consistent with the smoothed proportions, the following procedures were followed:

1. For each booklet, respondents were classified according to smoothing cell, and the raw population proportions for each cell were obtained. For example, the raw proportions for a booklet given to 9-year-olds in smoothing cell 4 was the total of the weights of all 9-year-olds receiving the booklet who were black and in the Southeastern region, divided by the total of the weights of all respondents receiving the booklet.

2. For each booklet and smoothing cell, a weight adjustment factor as the ratio of the smoothed population proportion (for the appropriate age, year and smoothing cell) over

TABLE E-4. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 17-Year-Olds

	Cell 1	2	3	4	5	6
Race	White	White	White	Black	Black	Other
Region	All	All	All	SE+	Not SE	All
Size of Community	BC+FR#	MC++	SP##	All	All	All
Year						
1969-70	0.3405	0.1447	0.3686	0.0415	0.0581	0.0466
1970-71	0.3419	0.1386	0.3704	0.0427	0.0591	0.0472
1971-72	0.3432	0.1326	0.3722	0.0440	0.0602	0.0478
1972-73	0.3446	0.1265	0.3740	0.0452	0.0612	0.0484
1973-74	0.3460	0.1204	0.3759	0.0465	0.0622	0.0491
1974-75	0.3474	0.1143	0.3777	0.0477	0.0633	0.0497
1975-76	0.3487	0.1082	0.3795	0.0490	0.0643	0.0503
1976-77	0.3501	0.1021	0.3813	0.0502	0.0653	0.0509
1977-78	0.3515	0.0961	0.3831	0.0515	0.0664	0.0515
1978-79	0.3529	0.0900	0.3849	0.0527	0.0674	0.0522

+SE = Southeast.
#BC+FR = Big Cities + Fringes.
++MC = Medium Cities.
##SP = Small Places.

---

the raw population proportion was obtained.

3. The adjusted weights for an individual were the product of that individual's original weight and the appropriate adjustment factor.

Changes in Smoothed Proportions as New Assessments Are Completed

Every time an assessment is completed, a new time point is added to each of the sequences of population proportions within the smoothing cells. This means that, even though robust/resistant procedures are used, the addition of a new point may somewhat change the values of smoothed proportions for prior years. Additionally, any changes in methodology will impact the estimates.

This means that the smoothed proportions, with the addition of the next assessment data, are apt to differ somewhat from the corresponding smoothed proportions without the new data. National Assessment has adopted the philosophy that the smoothed proportions, based on all currently available data using the best available algorithm, are the best available. Therefore, all subsequent analyses, for any year, will be done using this best available information, even though this may produce estimates that slightly differ from prior values.

GLOSSARY OF NATIONAL ASSESSMENT TERMS

Acceptable response. Any response to an exercise that demonstrates achievement of the objective measured by that exercise.

Administration time. The total time allowed on the paced audio tape for an exercise. (Includes the time allowed for the stimulus and the response.)

Administration timetable. Time periods during the school year when the various age groups are assessed. The time periods are:

October-December	13-year-olds
January-February	9-year-olds
March-May	17-year-olds

Age group or age level. Three age groups have been sampled in both music assessments: 9-year-olds, 13-year-olds and 17-year-olds attending school. Birth date ranges for each age group in each of the two assessments are as follows:

Assessment	Age 9	Age 13	Age 17
1971-72	1962	1958	10/54-9/55
1978-79	1969	1965	10/61-9/62

Assessment. The documentation of the progress in knowledge, skills and attitudes of American youth. Measures are taken at periodic intervals for each learning area, with the goal of determining trends and reporting the findings to the public and to the education community.

Assessment administrator. Individual employed to administer the assessment in participating schools.

Background questions. Questions about respondents' instructional experiences with music in school and out of school were included in music item booklets. Standard background questions asked in every learning area are found on the back pages of the item booklets and include such things as level of parental education and reading materials in the home. Background questions used in the 1978-79 music assessment appear in Appendix B.

Booklet. Items (exercises) are presented to respondents in booklets. Booklets are designed to be scored by optical scanning machines. Each booklet contains (1) instructions for answering items and sample items, (2) assessment items and (3) background questions. Each booklet con-

tains approximately 30-35 minutes of assessment items and 10-15 minutes of introductory material and background questions.

Category (scoring). A classification of a response to an open-ended exercise. See scoring guide.

Category within a variable. A subclassification within a variable. For example, male and female are categories of the variable sex. See reporting groups.

Difficulty level. The percentage of acceptable responses to an exercise.

Exercise. A task designed to measure an objective. Because NAEP does not administer "tests," but instead describes educational achievement over time, the term "exercise" is often used instead of the term "item" or "test item." The terms "item" and "exercise" are used synonymously in this report.

Exercise booklet. See booklet.

Exercise part. See item part.

Exercise pool. The entire set of exercises prepared for a learning area. This set includes recycled exercises, exercises developed for previous assessments but not used due to exercise booklet or budgetary constraints and newly developed exercises.

Field test. A pretest of exercises to obtain information regarding clarity, difficulty levels, timing, feasibility and special administrative problems needed for revision and selection of exercises to be used in the assessment.

Grade in school. Results are reported for 9-year-olds in the 3rd and 4th grades; 13-year-olds in the 7th and 8th grades; and 17-year-olds in the 10th, 11th and 12th grades.

Group administration. Booklets were administered to groups of 10 to 25 students in 1978-79. In 1971-72, group size was 12 respondents. A paced audio tape was used to provide uniform instructions and oral presentation of exercises.

Hand scoring (scoring). The coding of responses in a format compatible with the optical scanning equipment being used. Multiple-choice exercises can be directly machine scored; however, responses to open-ended exercises must be coded in scoring ovals so that they can then be machine scored. See scoring guide.

ID number. An identification number referring to the unique number assigned to each respondent. This number is assigned to preserve the anonymity of each respondent. NAEP does not keep records of the names of any individuals.

Item. See exercise.

Item booklet. See booklet.

Item part. Each part of an item that asks a separate question. Parts may all pertain to one stimulus, such as a graph or a table, or may concern the same topic.

Jackknife. The name of the algorithm used by NAEP to estimate standard errors of percentages and other statistics.

Learning area. One of the areas assessed by National Assessment: reading/literature, writing, mathematics, science, citizenship/social studies, art, music, career development. Also called "subject area."

Level of parental education. These levels are described in Appendix A.

Modal grade. The grade in which the majority of each in-school age group is enrolled. For 9-year-olds, the modal grade is the 4th grade; for 13-year-olds, the 8th grade; and for 17-year-olds, the 11th grade.

Objective. A desirable education goal agreed upon by scholars in the field, educators and concerned lay persons, and established through the consensus approach.

Objectives redevelopment. After the initial assessment of a learning area, one of the first steps in preparing for reassessment is a review of the learning-area objectives. This is carried out by scholars in the field, educators and concerned lay persons. These reviews may result in revision, modification or total rewriting of the learning-area objectives to reflect current curricular goals and emphases; they may also result in the endorsement of the objectives from the previous assessment as adequate for the next assessment.

Open-ended exercise. A nonmultiple-choice exercise that requires some type of written or oral response.

Paced audio tape. A tape recording that accompanies each booklet to ensure uniformity in administration. Instructions and exercises are played back from the tape recording so that reading difficulties will not interfere with an individual's ability to respond. Response time is included on the tape.

Primary sampling unit (PSU). First-stage sampling units, typically a county or a group of contiguous counties.

Principal's questionnaire. A data collection form given to school officials. The officials respond to questions concerning enrollments, size of the community, occupational composition of the community, and

so forth. Samples of these questionnaires are found in Appendix B. See also supplementary principal's questionnaire.

PSU. See primary sampling unit.

Public-use data tapes. Computer tapes containing respondent-level exercise and background/demographic data and machine-readable documentation. These tapes are available for use by external researchers wishing to do secondary analyses of National Assessment data.

Racial/ethnic category. For the music assessments, results are reported for whites, blacks and Hispanos.

Receipt control. Procedures implemented by scoring staff to check in and screen materials from the field. Information gained from receipt control procedures is relayed to assessment administrative staff so that any errors may be corrected.

Recycled exercises. The set of exercises that is kept secure from one assessment to the next to be used to measure changes (growth, stability or decline) in performance for the learning area.

Region. One of four geographical regions used in gathering and reporting data: Northeast, Southeast, Central and West. States included in each region are shown in Appendix A.

Released exercise. An exercise for which results and exercise text have been reported to the public.

Released exercise set. A set of released exercises, including documentation and scoring guides, that can be purchased from National Assessment. Data for the released exercise set are included as addendum pages.

Reporting groups. Categories of variables for which National Assessment data are reported. Variable categories are defined in Appendix A.

Rescore. If an open-ended exercise was scored under different conditions than presently held or if passage of time might affect scoring, responses from a previous assessment may be rescored at the same time that responses from a later assessment are scored. Responses from an earlier assessment also may be held and not scored so that they can be scored with responses from a later assessment.

Respondent. A person who responds to the exercises in an assessment booklet.

Response options. Different alternatives to a multiple-choice question that can be selected by the respondent.

Review conference. A conference held to review the objectives of a learning area to assure their acceptance as measures of the objectives by scholars, educators and lay persons or to review exercises for racial, ethnic, social or regional bias.

Sample. National Assessment does not assess an entire age population but rather selects a representative sample from the age group to answer assessment items. (See Chapter 4 for a description of National Assessment sampling procedures.)

Scoring guide. A guide for hand scoring an open-ended exercise that specifies descriptive or diagnostic categories by giving definitions and sample responses.

Scoring ovals. Scannable ovals printed beside multiple-choice options and printed at the bottom of the page for open-ended exercises (to be used in hand scoring). When ovals are marked, they can be scored by machine and responses recorded by computer.

Sex. Results are reported for males and females.

Size of community. Results are reported for four size-of-community categories: big cities, fringes around big cities, medium cities and small places. These categories are defined in Appendix A.

SMSA. Standard Metropolitan Statistical Area. SMSAs are economic and social units defined by the U.S. Bureau of the Census.

Standard error. A measure of sampling variability for a statistic. Because of NAEP's complex sample design, standard errors are estimated by jackknifing first-stage sample estimates.

Stem. The portion of an exercise that states the problem or asks the question.

Stimulus. For music exercises, this may be an aural or a visual stimulus used as part of the stem.

Subject area. See learning area.

Subpopulation or subgroup. Groups within the national population, such as males and females, for which data are reported.

Supplementary principal's questionnaire. A data collection form given to school officials. On this form, officials are asked to respond to questions concerning course offerings, materials and staffing specific to the learning area being assessed. A sample of this questionnaire is found in Appendix B. See also principal's questionnaire.

Tapescript. A script prepared for the announcer to use in producing the paced tape. It indicates exactly what is to be read or not read aloud to the students and indicates the amount of response time allowed for each exercise. See paced audio tape.

Timing. Most NAEP exercises are administered with a paced audio tape to standardize data collection conditions. The tape includes the amount of time students are allowed to respond to each exercise.

Type of community. Results are reported for three type-of-community categories: disadvantaged urban, advantaged urban and rural. Definitions of these categories are found in Appendix A.

User tape. See public-use data tape.

Variable. A classification of respondents. Standard reporting variables are: region, sex, race/ethnicity, level of parental education, size of community, type of community and grade in school.

Weight. A multiplicative factor equal to the reciprocal of the probability of a respondent being selected for assessment with adjustment for nonresponse -- an estimate of the number of persons in the population represented by a respondent in the sample. Theoretically, the sum of weights for all respondents at an age level is equal to the number of persons in the country at that age level.

BIBLIOGRAPHY

A Perspective on the First Music Assessment, Report no. 03-MU-02, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1974. ERIC no. ED 097 276. ISBN 0-89398-178-8.

An Assessment of Attitudes Toward Music, Report no. 03-MU-03, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1974. ERIC no. ED 099 270. ISBN 0-89398-179-6.

Chromy, J. and D. Horvitz. "Structure of Sampling and Weighting," 1969-70 Science: National Results and Illustrations of Group Comparisons, Report no. 1, 1969-70 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1970. ERIC no. ED 055 786. ISBN 0-89398-276-8.

Final Report on National Assessment of Educational Progress Sampling and Weighting Activities for Assessment Year 10. Research Triangle Park, N.C.: Research Triangle Institute, 1980.

Final Report on Year 10 Field Operations and Data Collection Activities, National Assessment of Educational Progress. Research Triangle Park, N.C.: Research Triangle Institute, 1979.

Final Report on Year 10 In-School Field Operations and Data Collection Activities, National Assessment of Educational Progress. Research Triangle Park, N.C.: Research Triangle Institute, 1979.

Folsom, Ralph E. National Assessment Approach to Sampling Error Estimation, Sampling Error Monograph 25U-796-5. Research Triangle Park, N.C.: Research Triangle Institute, 1977.

Kalsbeek, W.D. et al. No Show Analysis, final report. Raleigh, N.C.: Research Triangle Institute, 1975.

Miller, R.G. Jr. "A Trustworthy Jackknife," Annals of Mathematical Statistics, no. 35, 1964.

Miller, R.G. Jr. "Jackknifing Variances," Annals of Mathematical Statistics, no. 39, 1968.

Miller, R.G. Jr. "The Jackknife -- A Review," Biometrika, no. 61, 1974.

Moore, R. et al. The National Assessment Approach to Sampling. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1974. ERIC no. ED 099 416.

Mosteller, F. and J.W. Tukey. Data Analysis and Regression. Reading, Mass.: Addison-Wesley, 1977.

Mosteller, F. and J.W. Tukey, "Data Analysis Including Statistics," in Handbook of Social Psychology, 2nd ed., edited by E. Aronson and G. Lindzey. Reading, Mass.: Addison-Wesley, 1968.

Music 1971-79: Results From the Second Music Assessment, Report no. 10-MU-01, 1971-72 and 1978-79 Assessments. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1981. ISBN 0-89398-189-3.

Music Objectives, Report no. 03-MU-10, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1970. ERIC no. ED 063 197. ISBN 0-89398-175-3.

Music Objectives, Second Assessment, no. 10-MU-10, 1978-79 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1980. ISBN 0-89398-186-9.

Music Technical Report: Exercise Volume, Report no. 03-MU-20, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1975. ERIC no. ED 120 086. ISBN 0-89398-182-6.

Music Technical Report: Summary Volume, Report no. 03-MU-21, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1975. ERIC no. ED 114 348. ISBN 0-89398-183-4.

Rogers, W.T. et al. "Assessment of Nonresponse Bias in Sample Surveys: An Example From National Assessment," Journal of Educational Measurement, vol. 14, no. 4, 1977.

The First Music Assessment: An Overview, Report no. 03-MU-00, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1974. ERIC no. ED 097 275. ISBN 0-89398-181-8.

The First National Assessment of Musical Performance, Report no. 03-MU-01, 1971-72 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1974. ERIC no. ED 155 126. ISBN 0-89398-176-1.

The Second Assessment of Music, 1978-79, Released Exercise Set, no. 10-MU-25. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1980. ISBN 0-89398-187-7.

Tukey, John W. Exploratory Data Analysis. Reading, Mass.: Addison-Wesley, 1977.

Tukey, John W. "Some Considerations on Locators Apt for Some Squeezed-Tail (and Stretched-Tail) Parents" (paper produced in connection with research at Princeton University, supported by the Army Research Office), summer 1975.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1981—576-815/67 Region No. 8

(Continued From Inside Front Cover)

ART

1st Assessment (1974-75)

06-A-01	Design and Drawing Skills, June 1977	3.70
06-A-02	Knowledge About Art, January 1978	2.15
06-A-03	Attitudes Toward Art, May 1978	2.65
06-A-20	Art Technical Report: Exercise Volume, January 1978	27.50
06-A-21	Art Technical Report: Summary Volume, June 1978	6.80

2nd Assessment (1978-79)

10-A-01	Art and Young Americans, 1974-79: Results From the Second National Art Assessment , December 1981	8.90
10-A-25	The Second Assessment of Art, 1978-79, Released Exercise Set, April 1980	13.65

BACKGROUND REPORT

03/04-GIY	General Information Yearbook. A condensed description of the Assessment's methodology, December 1974	2.75

In addition to the above reports, National Assessment has produced reports in the areas of social studies, citizenship, writing, literature, reading, mathematics, science and career and occupational development. A complete publications list and ordering information are available from the address below.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
Suite 700, 1860 Lincoln Street
Denver, Colorado 80295